

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 11:57:59 ; Search time 57.2353 Seconds

(without alignments)  
103.669 Million cell updates/sec

Title: US-09-509-449C-5

Perfect score: 123

Sequence: 1 PROSRPSPMGCTDPRHRRNNG 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 645147

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003s:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	100.0	21	2	AAW99648 Hepatitis
2	123	100.0	21	3	AAW99648 Hepatitis
3	112	91.1	24	2	AAW99648 Hepatitis
4	112	91.1	25	6	AAW99648 Hepatitis
5	112	91.1	28	2	AAW99648 Hepatitis
6	112	91.1	29	2	AAW99648 Hepatitis
7	111	90.2	20	2	AAW99648 Hepatitis
8	110	89.4	20	2	AAW99648 Hepatitis
9	110	89.4	20	2	AAW99648 Hepatitis
10	105	85.4	20	2	AAW99648 Hepatitis
11	105	85.4	20	2	AAW99648 Hepatitis
12	105	85.4	20	2	AAW99648 Hepatitis
13	105	85.4	20	2	AAW99648 Hepatitis
14	100	81.3	27	2	AAW99648 Hepatitis
15	94	76.4	18	5	AAW99648 Hepatitis
16	94	76.4	19	5	AAW99648 Hepatitis
17	94	76.4	20	2	AAW99648 Hepatitis
18	90	73.2	30	5	AAW99648 Hepatitis
19	85	69.1	15	2	AAW99648 Hepatitis
20	83	67.5	15	4	AAW99648 Hepatitis
21	83	67.5	15	4	AAW99648 Hepatitis
22	83	67.5	15	4	AAW99648 Hepatitis
23	81	65.9	15	4	AAW99648 Hepatitis
24	78	63.4	15	6	AAW99648 Hepatitis
25	67	54.5	15	7	AAW99648 Hepatitis

26	67	54.5	16	4	AAJ03982 Hepatitis
27	67	54.5	20	2	AAW99648 Hepatitis
28	65	52.8	15	7	AAW99648 Hepatitis
29	64	52.0	31	2	AAW99648 Hepatitis
30	62	50.4	15	4	AAW99648 Hepatitis
31	61	49.6	11	4	AAW99648 Hepatitis
32	61	49.6	11	4	AAW99648 Hepatitis
33	57	46.3	18	3	AAW99648 Hepatitis
34	57	46.3	20	2	AAW99648 Hepatitis
35	57	46.3	25	2	AAW99648 Hepatitis
36	56	45.5	20	2	AAW99648 Hepatitis
37	56	45.5	20	3	AAW99648 Hepatitis
38	52	42.3	9	4	AAW99648 Hepatitis
39	52	42.3	9	4	AAW99648 Hepatitis
40	52	42.3	9	4	AAW99648 Hepatitis
41	52	42.3	9	4	AAW99648 Hepatitis
42	52	42.3	9	4	AAW99648 Hepatitis
43	51	41.5	11	4	AAW99648 Hepatitis
44	51	41.5	11	4	AAW99648 Hepatitis
45	51	41.5	11	4	AAW99648 Hepatitis

## ALIGNMENTS

### RESULT 1

ID AAW99648 standard; peptide; 21 AA.

AC AAW99648;  
DT 21-MAY-1999 (first entry)  
DE Hepatitis C virus peptide SEQ ID NO:5.

KW Hepatitis C virus; HCV; hepatitis B virus; diagnosis;  
KW detection viral infection.  
XX Hepatitis C virus.

OS Hepatitis C virus.

PN WO9906836-A1.

PD 11-FEB-1999.

PF 04-AUG-1998; 98WO-JP003476.

PR 04-AUG-1997; 97JP-00209515.

PR 04-AUG-1997; 97JP-00209522.

PR 31-UTL-1998; 98UP-00218136.

PA (TOFU) TONEN CORP.

PI Aoyagi K, Ohue C, Iida K, Kimura T, Yagi S;

DR WPI, 1999-153953/13.

PT Detection and assay for, e.g. hepatitis C virus - comprises treatment of

PT the sample with surfactants; protein denaturing agent prior to

PS immunosassay.

PS Example 3; Page 88; 98pp; Japanese.

CC A method has been developed for virus-containing sample is to use a  
CC treatment solution comprising: (i) an anionic surfactant; (ii) a protein  
CC amphiphilic surfactant; (iii) a non-ionic surfactant; and (iv) a protein  
CC denaturing agent. Also described in the present invention are: (1) a  
CC similar treatment method by using a solution containing: (i) a chaotropic  
CC ion; (ii) an acidifying agent; and (iii) a non-ionic surfactant; (2) an  
CC assaying method for viruses by quantitative detection of the presence of  
CC viral antigen using a probe for reaction with the specifically recognised  
CC viral antigen; (3) a hybridoma cell line chosen from HC11-11 (FERM BP-  
CC 6005), HC11-14 (FERM BP-6006), HC11-10 (FERM BP-6004), HC11-3 (FERM BP-  
CC 6002) and HC11-7 (FERM BP-6003); (4) a monoclonal antibody produced by

CC the hybridoma; (5) a quantitation kit or diagnostic reagent for  
CC determining the presence of a virus in the sample, especially hepatitis C  
CC virus (HCV), using the immunoassay using: (i) a treatment solution  
CC containing an anionic surfactant or chaotropic agent, or urea, imidazole  
CC ring-bearing or indole ring-bearing compound(s); and (ii) the monoclonal  
CC antibody; and (6) assaying for viruses by detecting, e.g., viral antigen  
CC (viral antibody) in the presence of a surfactant containing not less than  
CC 10C  $\alpha$ 1y1 and a secondary, tertiary or quaternary amine groups and/or a  
CC non-ionic surfactant with hydrophilic-lipophile balance (HLB) of 12-14,  
CC during which the viral antigen binds with its probe. The methods are  
CC useful in detecting or assaying viruses e.g., HCV and hepatitis B virus,  
CC in blood and its preparations and in diagnosis of viral infections e.g.,  
CC in humans. These methods are highly sensitive and specific, simple to  
CC operate, giving quantitative results. The present sequence represents an  
CC HCV peptide from the present invention  
CC  
CC  
CC Sequence 21 AA;  
CC  
CC

Query Match	100.0%;	Score 123;	DB 2;	length 21;
Best Local Similarity	100.0%;	Pred. No. 3.2e-10;		
Matches 21; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	PRGSRPSWGPDPHRSRNVG	21
Db	1	PRGSRP SWGPDPHRSRNVG	21

RESULT 2	
AAy80190	
ID	AAy80190 standard; peptide; 21 AA

AC	AAV80190;
XX	
DT	24-MAY-2000 (first entry)

Hepatitis C virus peptide sequence SEQ ID NO:5.

KW Hepatitis C virus; HCV; antigen; antibody; infection; blood donation;  
KM screening.

OS Hepatitis C virus.

PN WO200007023-A1.

PD 10-FEB-2000.

PF 30-JUL-1999; 99WO-JP004129.

PR 30-JUL-1998; 98JP-00216094.

PA (ADLI-) ADVANCED LIFE SCI INST INC.

PI Aoyagi K, Ohue C, Iida K, Yagi S;

WPI; 2000-195352/17.

PT Method of assaying for hepatitis C virus infection comprises binding  
PT viral core antigen or anti-hepatitis C virus antibodies with a probe in  
PT the presence of a surfactant.

PS Example 3; Page 45; 55pp; Japanese.

CC A method has been developed of assaying for hepatitis C virus (HCV). The  
CC method comprises binding a viral core antigen or anti-HCV antibody to a  
CC suitable prime in the presence of a non-ionic surfactant and/or a  
CC cationic surfactant with secondary, tertiary or quaternary amine groups  
CC and carbon chains with 10 or more carbon atoms. The assay is useful for  
CC the detection of HCV infection in biological samples, especially in  
CC donated blood. The method is highly sensitive for detecting HCV and  
CC reduces the window period between negative and positive screening tests  
CC for HCV on blood donated by a recently infected person. The present  
CC sequence represents an HCV peptide which is used in an example from the  
CC present invention

XX Sequence 21 AA;  
SQ

Query Match	100.0%;	Score 123;	DB 3;	Length 21;
Best Local Similarity	100.0%;	Pred. No. 3.2e-10;		
Matches 21; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

```

Qy      1 PRGSRPSWGPTDPRHRSKRVG 21
         |||||
Db      1 PRGSRPSWGPTDPRHRSKRVG 21

```

RESULT 3  
AAR34055  
ID AAR34055 standard; peptide; 24 AA.

DT 25-MAR-2003 (revised)  
DT 22-JUL-1993 (first entry)  
XX  
DE HCV C domain peptide DP10.

KM	Hepatitis C virus; non structural region; antigen; diagnosis; vaccine.
XX	
OS	Synthetic.

PN WO9306488-A1.

PD 01-APR-1993.

PF 16-SEP-1992; 92WO-US007865.

PR 16-SEP-1991; 91US-00762135

XX  
P2 (CENT-1) GENERALS TECHNOLOGIES INC

XX	Dreesman GB	Burk KH	Paulletti D
PI			

XX  
DB WBT: 1993-117737/14.

XX  
PT    Detection of hepatitis C virus antigens - using HCV reactive antibody

bound to solid support and competitive HCV antigen-reporter complex.

aa Example 7; Page 130; 190pp; English.  
PS

The synthetic peptide corresponds to residues 98-121 of the C domain of HCV. The peptide may be used for the sensitive and specific detection of HCV infection. Antibodies raised against the peptide can be used for passive immuno-prophylaxis and the HCV antigens can be used in vaccines to prevent HCV infection. See also AAR34041-65. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 24 AA;

Query Match	91.1%;	Score 112;	DB 2;	Length 24;
Best Local Similarity	90.5%;	Pred. No. 1.2e-08;		
Matches 19;	Conservative	1;	Mismatches	1; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRHRSRNVG 21  
|||||:|  
Db 3 PRGSRPSWGPTDPRRSRLTG 23

RESULT 4  
ABJ37517  
ID ABJ37517 standard; peptide; 25 AA.

AA ABJ37517;  
AC

DT	10-MAY-2003	(first entry)
----	-------------	---------------

XX

```

DE HCV-core derived peptide 1.
XX
XX Hepatitis C virus; HCV; antigen; antibody.
XX
XX Hepatitis C virus.
XX
XX WO2003002749-A2.
XX
XX
XX 09-JAN-2003.
XX
XX 24-JUN-2002; 2002WO-US019958.
XX
XX 26-JUN-2001; 2001US-00891983.
XX
XX 17-JUN-2002; 2002US-00173480.
XX
XX
XX (ABBOTT LAB.
XX
XX Shah DO, Dawson GA, Muerhoff AS, Jiang L, Gutierrez RA, Leary TP,
XX
XX Desai S, Stewart JL;
XX
XX WPI; 2003-201507/19.
XX
XX
XX Detecting hepatitis C virus antigen and antibody in sample, by contacting
XX
XX sample with HCV antigen and antibody, for forming HCV antigen-test
XX
XX antibody complex and HCV antibody-test antigen complex, detecting
XX
XX complexes.
XX
XX
XX Example 1; Page 17; 92pp; English.
XX
XX
XX The invention relates to a novel method for simultaneously detecting a
XX
XX hepatitis C virus (HCV) antigen and an HCV antibody in a test sample. The
XX
XX novel method comprises contacting the test sample with the antigen and
XX
XX antibody, both coated on solid phases, independently, for a time and
XX
XX under conditions sufficient for the formation of an antibody/antigen
XX
XX complex and a second antibody/antigen complex, and detecting the two
XX
XX complexes, which indicate the presence of the antigen and antibody in the
XX
XX test sample. The method is useful for detecting a hepatitis C virus (HCV)
XX
XX antigen and HCV antibody in a test sample. This sequence represents an
XX
XX HCV-core derived peptide relating to the HCV antigen/antibody detection
XX
XX method of the invention
XX
XX
XX Sequence 25 AA;
XX
XX
XX Query Match 91.1%; Score 112; DB 6; Length 25;
XX
XX Best Local Similarity 90.5%; Pred. No. 1.3e-08;
XX
XX Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX
XX 1 PRGSRPSPGPTDPPRRSRNVG 21
XX
XX 5 PRGSRPSPGPTDPPRRSRNVG 25
XX
XX
XX RESULT 5
XX
XX AAR33863
XX
XX AAR33663 standard; peptide; 28 AA.
XX
XX
XX AAR33863;
XX
XX
XX 25-MAR-2003 (revised)
XX
XX 19-JUL-1993 (first entry)
XX
XX
XX Polypeptide p99 comprising HCV viral antigen.
XX
XX
XX Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG;
XX
XX p408.
XX
XX Synthetic.
XX
XX
XX WO9306247-A1.
XX
XX
XX 01-APR-1993.
XX
XX
XX 16-SEP-1992; 92WO-US00781.3.
XX
XX

```

XX 16-SEP-1991; 91US-00760292.  
PR  
XX  
XX (ABBOTT ) ABBOTT LAB.  
XX  
XX Leelewski R, Leung TK;  
XX  
XX WPI; 1993-117563/14.  
DR  
XX  
XX Assay for detecting presence of antibody to hepatitis C viral antigen -  
PT by contacting sample with polypeptide contg. at least one epitope of  
PT virus antigen.  
PT  
XX  
XX Disclosure, Page 12; 63pp; English.  
PS  
XX  
XX The synthetic peptide p99 represents amino acid residues 99-126 of the  
CC hepatitis C viral antigen. The peptide may be used in an assay to detect  
CC antibodies to HCV and thus to diagnose chronic HCV infection. See also  
CC AR33862-87. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 28 AA;  
SQ

CY	1	PROSRPSWGPTDPRHRSRVG	21	91.1%;	Score 112;	DB 2;	Length 28;
				90.5%;	Pred. No. 1.4e-08;		
Db	2	PROSRPSWGPTDPRHRSRVG	22	1;	Mismatches 1;	Indels 0;	Gaps 0

  

RESULT 6	
ID	AAR13347
	AAR13347 standard; protein; 29 AA.
XX	
AC	AAR13347;
XX	
DT	25-MAR-2003 (revised)
DT	23-OCT-1991 (first entry)
DE	P99 HCV antigen (99-126).
XX	
KM	CI100-3; hepatitis C virus; immunoassay; epitope.
XX	
OS	Synthetic.
PN	AU9068390-A.
XX	
PD	27-JUN-1991.
XX	
PF	21-DEC-1990; 90AU-00068390.
XX	
ER	22-DEC-1989; 89US-00456162.
ER	07-NOV-1990; 90US-00610180.
XX	
PA	(AB90) ABBOTT LAB.
PA	(LESN/) LESNIEWSKI R. R.
XX	
WF1	1991-238393/33.
XX	
PT	Immunological assays for hepatitis C virus antibody - by using
PT	polypeptide(s) contg. epitope(s) of hepatitis C virus antigens.
XX	
PS	Claim 10; Page 48; 62pp; English.
XX	
CC	The polypeptide may be prepared by solid phase synthesis fragment
CC	coupling (pret.) or using recombinant technology. The assay has increased
CC	sensitivity and is more specific than assays using the polypeptide CI100-3
CC	(EP-118216). See also AA013146-48 and AAR13343-65. (Updated on 25-MAR-
CC	2003 to correct PA field.)
XX	
XX	
Sequence 29 AA;	

Query Match 91.1%; Score 112; DB 2; Length 29;  
 Best Local Similarity 90.5%; Pred. No. 1.5e-08;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RGSRPWGPTDPRHRSRVNG 21  
 |||||  
 DB 3 RGSRPWGPTDPRHRSRVNG 23

## RESULT 7

AA38293  
 ID AA38293 standard; protein; 20 AA.

AA38293;

DT 21-OCT-1993 (first entry)

DE NAMB hepatitis virus HC-OM gene polypeptide CP-5-2.

KW Non-A, non-B; virus; polymerase chain reaction; detection; sensitive;  
 KM specific; HCV; NANBH.

OS Non-A.  
 non-B hepatitis virus.

FN JP05091884-A.

PD 16-APR-1993.

PE 10-APR-1991; 91JP-00196175.

PR 12-JUN-1990; 90JP-00153401.

PR 08-NOV-1990; 90JP-00304405.

PA (NAXA/) NAKAMURA T.

DR WPI; 1993-199637/25.

DR N-PSDB; AAQ43903.

PT Antigen related to non-A and non-B hepatitis virus - comprises non-  
 translation region comprising 340 - 341 moles. of nucleotides, non-  
 translation region comprising 1885 - 2551 moles. of nucleotides including  
 region 1,149 and, etc.

PS Claim 29; Page 63; 73pp; Japanese.

CC The sequence is that of NAMB hepatitis virus HC-OM gene polypeptide CP-5-  
 2. It may be used in a system for detecting NAMB hepatitis. This  
 method is highly specific and sensitive, and can detect NAMB hepatitis  
 virus which could not be detected by conventional methods

SO Sequence 20 AA;

Query Match 90.2%; Score 111; DB 2; Length 20;

Best Local Similarity 95.0%; Pred. No. 1.4e-08;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RGSRPWGPTDPRHRSRVNG 21  
 |||||  
 DB 1 RGSRPWGPTDPRHRSRVNG 20

## RESULT 8

AA38294  
 ID AA38294 standard; protein; 20 AA.

AA38294;

DT 21-OCT-1993 (first entry)

DE NAMB hepatitis virus HC-OM gene polypeptide CP-5-3.

KW Non-A, non-B; virus; polymerase chain reaction; detection; sensitive;

KW specific; HCV; NANBH.  
 XX Non-A.  
 OS non-B hepatitis virus.

PN JP05091884-A.

PD 16-APR-1993.

PE 10-APR-1991; 91JP-00196175.

PR 12-JUN-1990; 90JP-00153401.

PR 08-NOV-1990; 90JP-00304405.

PA (NAXA/) NAKAMURA T.

DR WPI; 1993-199637/25.

DR N-PSDB; AAQ43904.

PT Antigen related to non-A and non-B hepatitis virus - comprises non-  
 translation region comprising 340 - 341 moles. of nucleotides, non-  
 translation region comprising 1885 - 2551 moles. of nucleotides including  
 region 1,149 and, etc.

PS Claim 30; Page 63; 73pp; Japanese.

CC The sequence is that of NAMB hepatitis virus HC-OM gene polypeptide CP-5-  
 3. It may be used in a system for detecting NAMB hepatitis. This  
 method is highly specific and sensitive, and can detect NAMB hepatitis  
 virus which could not be detected by conventional methods

SO Sequence 20 AA;

Query Match 89.4%; Score 110; DB 2; Length 20;

Best Local Similarity 90.0%; Pred. No. 2e-08; 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGSRPWGPTDPRHRSRVNG 21  
 |||||  
 DB 1 RGSRPWGPTDPRHRSRVNG 20

AA67636  
 ID AA67636 standard; protein; 20 AA.

AA67636;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 18-AUG-1995 (first entry)

DE Epitope of Non-A Non-B hepatitis virus structural protein.

KW Non-A Non-B hepatitis virus; structural region; cDNA to genomic RNA;  
 detection; reagent; anti-Non-A Non-B hepatitis virus antibody; vaccine;  
 antigen; epitope; diagnosis.

OS Non-A.  
 non-B hepatitis virus.

PN EP628572-A2.

PD 14-DEC-1994.

PE 27-MAY-1994; 94EP-00108256.

PR 28-MAY-1993; 93JP-00126709.

PR 02-MAR-1994; 94JP-00032201.

PA (EISA) EISAI CO LTD.

PI Aoyama M, Obara T, Tohmatu J, Sawada T, Hosoda T, Iwasaki Y;

PI Arima T;  
XX  
DR WPI; 1995-015655/03.  
XX  
PT New non-A non-B hepatitis virus sub-type - used to develop prods. for  
PT detection, diagnosis, prevention and treatment of non-A non-B hepatitis.  
XX  
PS Example 3; Page 40; 59pp; English.  
XX  
CC AAR67636 is designated M35 and is located at positions 101-120 of the  
CC structural protein (AAR67617) of the Non-A Non-B (NANB) hepatitis virus  
CC encoded by a partial cDNA to genomic RNA sequence. Regions of the  
CC polypeptide were studied for suitability as an epitope. The positive  
CC ratios of the peptide epitope to the plasma of donors, who had been  
CC judged as positive with a reagent of the second generation, i.e., a  
CC second generational anti hepatitis virus antibody detection reagent, were  
CC determined. This peptide had a positive ratio (%) of 31.5. The novel  
CC epitopes are effective in the diagnosis of NANB hepatitis. The nucleotide  
CC sequences (see AAQ75817-19) were isolated from the plasma of donors in  
CC Japan with high s-GTP levels, and were found to be different from  
CC previously reported NANB hepatitis viruses. The DNA can be used as a  
CC reagent for detecting the NANB hepatitis viral gene. The polypeptides can  
CC be used as reagents for detecting anti-NANB hepatitis antibodies or as a  
CC NANB hepatitis viral vaccine. (Updated on 25-MAR-2003 to correct PN  
CC field.) (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 20 AA;  
XX  
Query Match 89.4%; Score 110; DB 2; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2e-08;  
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 2 RGSRRPSWGPTDPRRHSRVNG 21  
1 RGSRRPSWGPTDPRRHSRVNG 20  
XX  
RESULT 10  
AAR25132  
ID AAR25132 standard; protein; 20 AA.  
XX  
AC AAR25132;  
XX  
DT 23-DEC-1992 (first entry)  
XX  
DE mokk-C3.  
XX  
XX Hepatitis C virus; blood transfusion.  
XX  
OS Synthetic.  
XX  
PN JP04159298-A.  
XX  
PD 02-JUN-1992.  
XX  
PF 19-OCT-1990; 90JP-00282431.  
XX  
PR 19-OCT-1990; 90JP-00282431.  
XX  
PA (OLYU) OLYMPUS OPTICAL CO LTD.  
XX  
DR WPI; 1992-231947/28.  
XX  
PT New peptides acting as antigenic analogues of human hepatitis C virus -  
PT useful for detecting HCV antibody positive patients.  
XX  
PS Claim 1; Page 1; 14pp; Japanese.  
XX  
CC The sequences given in AAR25130-35 are peptides which have reactivity to  
CC the antibody against hepatitis C virus (HCV). They can be used on their  
CC own or as a mixture two different peptides. Using these peptides, HCV  
CC antibody positive patients can be detected and hepatitis caused by blood  
CC transfusion can be prevented

XX  
SQ Sequence 20 AA;  
XX  
Query Match 85.4%; Score 105; DB 2; Length 20;  
Best Local Similarity 90.0%; Pred. No. 9.9e-08;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
XX  
QY 2 RGSRRPSWGPTDPRRHSRVNG 21  
1 RGSRRPSWGPTDPRRHSRVNG 20  
XX  
RESULT 11  
AAR38292  
ID AAR38292 standard; protein; 20 AA.  
XX  
AC AAR38292;  
XX  
DT 21-OCT-1993 (first entry)  
XX  
DE NANB hepatitis virus HC-OM gene polypeptide CP-5-1.  
XX  
KW Non-A, non-B; virus; polymerase chain reaction; detection; sensitive;  
KW specific; HCV; NANBH.  
XX  
OS Non-A.  
OS non-B hepatitis virus.  
XX  
XX  
XX PN JP05091884-A.  
XX  
PD 16-APR-1993.  
XX  
PF 10-APR-1991; 91JP-00196175.  
XX  
PR 12-JUN-1990; 90JP-00153401.  
XX  
PR 08-NOV-1990; 90JP-00304405.  
XX  
PA (NAKA/) NAKAMURA T.  
XX  
DR WPI; 1993-199637/25.  
XX  
DR N-PSDB; AAQ43902.  
XX  
PT Antigen related to non-A and non-B hepatitis virus - comprises non-  
PT translation region comprising 340 - 341 mols. of nucleotides, non-  
PT translation region comprising 1885 - 2551 mols. of nucleotides including  
PT region 1,149 and, etc.  
XX  
XX  
PS Claim 28; Page 63; 73pp; Japanese.  
XX  
CC The sequence is that of NANB hepatitis virus HC-OM gene polypeptide CP-5-  
CC 1. It may be used in a system for detecting NANB hepatitis. This  
CC method is highly specific and sensitive, and can detect NANB hepatitis  
CC virus which could not be detected by conventional methods  
XX  
SQ Sequence 20 AA;  
XX  
Query Match 85.4%; Score 105; DB 2; Length 20;  
Best Local Similarity 90.0%; Pred. No. 9.9e-08;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
XX  
QY 2 RGSRRPSWGPTDPRRHSRVNG 21  
1 RGSRRPSWGPTDPRRHSRVNG 20  
XX  
RESULT 12  
AAR56608  
ID AAR56608 standard; protein; 20 AA.  
XX  
AC AAR56608;  
XX  
DT 24-MAR-1995 (first entry)  
XX

```

DE  HCV peptide NP-11.
XX
XX  Hepatitis C virus; HCV, core protein; epitope; T-cell; immunity; CD8;
KM  CDB.
XX
XX  Hepatitis C virus.
OS
XX  JPO6199894-A.
XX
XX  19-JUL-1994.
PD
XX
XX  02-APR-1993; 93JP-00076791.
XX
XX  27-AUG-1992; 92JP-00228965.
XX
XX  10-NOV-1992; 92JP-00299691.
XX
XX  (ASAH ) ASAH KASEI KOYO KK.
PA  (INOM/) INOMAMARI M.
XX
XX  WPI; 1994-269451/33.
DR
XX
XX  T cell epitope present in the core protein region of Hepatitis C virus
PT  (HCV) - used for activation of cellular immunity mechanisms.
XX
XX  Disclosure; Fig 2; 14pp; Japanese.
XX
XX  A T-cell stimulating peptide is claimed which is ca. 5-20 amino acids
CC  long and part of the core protein of HCV and is recognised by and
CC  stimulates T-cells. A CD8-positive T-cell stimulating peptide has the
CC  sequence given in AAR56606 (NP-9) or AAR56616 (9MA). A CD4-positive T-
CC  cell stimulating peptide has the sequence given in AAR56609 (NP-12) or
CC  AAR56614 (NP-17). Synthetic peptide mixts. (Mix A: NP-1 - NP-5, Mix B: NP
CC  -6 - NP-10, Mix C: NP-11 - NP14, MixD: NP-15 - NP-18) were used in
CC  -6 - NP-10, Mix C: NP-11 - NP14, MixD: NP-15 - NP-18) were used in
CC  are variants of peptide NP-9
CC
XX
SQ  Sequence 20 AA;

Query Match      85.4%; Score 105; DB 2; Length 20;
Best Local Similarity 90.0%; Pred. No. 9.9e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  2 RGSRRPSWGPDPDRRRSRNVG 21
    |||||
DB  1 RGSRRPSWGPDPDRRRSRNLG 20

RESULT 13
AAR56604
ID  AAR56604 standard; peptide; 30 AA.
XX
XX  AAR56604;
AC
XX
XX  08-MAY-2002 (first entry)
DT
XX
XX  HCV HepC1a segment 7.
DE
XX
XX  Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KM  viral infection; human immunodeficiency virus; melanoma;
KM  bacterial infection; Salmonella; Legionella; parasitic infection;
KM  Trypanosoma; Toxoplasma; Giardia.
XX
XX  Hepatitis C virus.
OS
XX
XX  WO200190197-A1.
XX
XX  29-NOV-2001.
PD
XX
XX  25-MAY-2001; 2001WO-AU000622.
XX
XX  26-MAY-2000; 2000AU-00007761.
XX
XX  (AUSU ) UNIV AUSTRALIAN NAT.

```

```

XX
XX  Thomson SA, Ramehaw IA;
PI
XX
XX  WPI; 2002-147575/19.
DR
XX
XX  N-P8DB; ABR36442.
DR
XX
XX  New synthetic polypeptides having several different segments of at least
PT  one parent polypeptide linked together differently compared to the
PT  linkage in the parent polypeptide, for inducing immune response against a
XX  pathogen or cancer.
XX
XX  Example 2; Fig 26; 364pp; English.
XX
XX
XX  The invention relates to a new synthetic polypeptide (I) comprising
CC  several different segments of at least one parent polypeptide linked
CC  together in a different relationship relative to their linkage in the
CC  parent polypeptide to impede, abrogate or otherwise alter at least one
CC  function associated with the parent polypeptide and for inducing an
CC  immune response against a pathogen or cancer. Also included are a
CC  synthetic polynucleotide encoding and a computer system for designing the
CC  synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC  are referred to as a Savine. The synthetic polypeptide is useful for
CC  modulating immune responses preferably directed against a pathogen or a
CC  cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC  and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC  oesophagus, brain, testicle, uterus), as potentiating agents.
CC  Compositions comprising the polypeptide may be used in the treatment or
CC  prophylaxis against viral (such as infections caused by HIV (human
CC  immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC  virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC  (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC  Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC  (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC  Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC  a peptide derived from a parent protein used to construct a savine of the
CC  invention
XX
XX
SQ  Sequence 30 AA;

Query Match      85.4%; Score 105; DB 5; Length 30;
Best Local Similarity 94.7%; Pred. No. 1.5e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 PRGSRPSWGPDPDRRRSRN 19
    |||||
DB  12 PRGSRPSWGPDPDRRRSRN 30

RESULT 14
AAR31922
ID  AAR31922 standard; peptide; 27 AA.
XX
XX  AAR31922;
AC
XX
XX  25-MAR-2003 (revised)
DT
XX
XX  03-MAR-1993 (first entry)
DT
XX
XX  HCV peptide BCH-443.
DE
XX
XX  Hepatitis C virus; vaccine; screening; diagnosis.
KM
XX
XX  Synthetic.
OS
XX
XX  EP507615-A1.
XX
XX  07-OCT-1992.
XX
XX  03-APR-1992; 92EP-00302975.
XX
XX  05-APR-1991; 91US-00681701.
XX
XX  (BIOC-) BIOCHEM PHARMA INC.
XX

```

PI Lacroix M;  
 XX  
 DR WPI; 1992-333826/41.  
 XX  
 PT Hepatitis C virus peptide(s) and their mixts. - used for diagnosing and  
 PT vaccinating against hepatitis C virus infections.  
 XX  
 PS Claim 1; Page 27; 34pp; English.  
 XX  
 CC The peptide BCH-443 was synthesised based upon amino acids 101-127 of  
 CC hepatitis C virus (HCV) envelope protein. It is useful as an active  
 CC ingredient in vaccines against HCV and can be used for screening and  
 CC diagnosing HCV infection. See also AAR31917-R31926. (Updated on 25-MAR-  
 CC 2003 to correct FN field.)  
 CC  
 XX  
 SQ Sequence 27 AA;  
 Query Match 81.3%; Score 100; DB 2; Length 27;  
 Best Local Similarity 85.0%; Pred. No. 6.6e-07;  
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RGSPPSWGPTDPRRRSRNVG 21  
 DB 1 RGSPPSWGPTDPRRRSRNVG 20

# RESULT 15

AAM50830  
 ID AAM50830 standard; peptide; 18 AA.

AC AAM50830;  
 XX  
 DT 07-AUG-2003 (revised)  
 DT 01-MAY-2002 (first entry)  
 XX

DE HCV core region peptide MDL-8 (aa103-120).

XX HCV; MDL-8; infection; diagnosis; vaccine.

XX Hepatitis C virus.

XX WO200204484-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-CA000988.

XX 07-JUL-2000; 2000CA-02311022.

XX (MEDM-) MEDMIRA INC.

XX Chan HKW, Theolis R;

XX WPI; 2002-179696/23.

PT New hepatitis C virus (HCV) peptides and mosaic antigen composition  
 PT comprising the peptides, useful as immunoreagents for detecting HCV  
 PT antibodies infection, or as an immunogen for stimulating production of  
 PT antibodies against HCV.

PS Example 1; Page 33; 59pp; English.

XX The present sequence is that of peptide MDL-8 comprising amino acid  
 CC residues 103-120 of the hepatitis C virus (HCV) polyprotein core region  
 CC (see AAM50816). A panel of overlapping peptides between positions 1 and  
 CC 186 of the HCV core region was synthesised and tested for  
 CC immunoreactivity with a commercially available human serum panel. All of  
 CC the immunoreactivity was concentrated within the core region between  
 CC amino acids 1 to 80. MDL-8 did not include a major immunodominant region.  
 CC The invention provides a highly immunoreactive mosaic antigen composition  
 CC (MAC) comprising 2 or more immunoreactive peptides (see AAM50816-27),  
 CC each of which is immobilised to a carrier. The unique combination of HCV  
 CC core peptides in the MAC provides higher specificity and sensitivity for

CC detection of human antibodies specificity to HCV in rapid HCV diagnostic  
 CC applications. An in vitro diagnostic method for detecting anti-HCV  
 CC antibodies in a test sample and a diagnostic test kit are provided which  
 CC use the MAC as an immunoreagent. The HCV MAC may also potentially be used  
 CC as an immunogen in vaccine compositions. (Updated on 07-AUG-2003 to  
 CC correct OS field.)  
 CC  
 XX  
 SQ Sequence 18 AA;

Query Match 76.4%; Score 94; DB 5; Length 18;  
 Best Local Similarity 88.9%; Pred. No. 3.1e-06;  
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 SRPSWGPTDPRRRSRNVG 21  
 DB 1 SRPSWGPTDPRRRSRNVG 18

Search completed: June 29, 2004, 12:05:56  
 Job time : 58.2353 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 12:03:34 ; Search time 17.2941 Seconds  
(without alignments)  
62.669 Million cell updates/sec

Title: US-09-509-449C-5  
Perfect score: 123  
Sequence: 1 PRGSRPSWGPTDPRHRSRVNG 21

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 227376

Minimum DB seq length: 0  
Maximum DB seq length: 40

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	100.0	21	US-10-133-007-5	Sequence 5, Appl1
2	112	91.1	24	PCT-US92-07865-15	Sequence 15, Appl1
3	112	91.1	28	US-08-905-054B-3	Sequence 3, Appl1
4	112	91.1	28	PCT-US92-07813-3	Sequence 3, Appl1
5	100	81.3	27	US-07-681-701-10	Sequence 10, Appl1
6	94	76.4	20	US-08-635-886C-8	Sequence 8, Appl1
7	94	76.4	20	US-08-974-690C-8	Sequence 8, Appl1
8	85	69.1	15	US-08-604-565-22	Sequence 22, Appl1
9	85	69.1	15	US-09-689-678-22	Sequence 22, Appl1
10	64	52.0	31	US-08-473-475A-25	Sequence 25, Appl1
11	57	46.3	20	US-08-635-886C-9	Sequence 9, Appl1
12	57	46.3	20	US-08-974-690C-9	Sequence 9, Appl1
13	56	45.5	20	US-10-133-007-6	Sequence 6, Appl1
14	50	40.7	8	US-08-444-818-301	Sequence 301, Appl1
15	49	39.8	8	US-08-444-818-299	Sequence 299, Appl1
16	49	39.8	8	US-08-444-818-302	Sequence 302, Appl1
17	48	39.0	8	US-08-444-818-300	Sequence 300, Appl1
18	44	35.8	29	US-09-205-258-330	Sequence 630, Appl1
19	43	35.0	31	US-09-308-345A-31	Sequence 31, Appl1
20	40.5	32.9	35	US-09-366-009-28	Sequence 28, Appl1
21	40.5	32.9	35	US-08-809-156B-28	Sequence 28, Appl1
22	39.5	32.1	40	US-09-314-268-106	Sequence 106, Appl1
23	39	31.7	31	US-09-071-224-36	Sequence 36, Appl1
24	38.5	31.3	40	US-09-314-268-107	Sequence 107, Appl1
25	38	30.9	8	US-08-444-818-298	Sequence 298, Appl1
26	38	30.9	31	US-09-308-345A-34	Sequence 34, Appl1
27	38	30.9	39	US-08-390-353A-11	Sequence 11, Appl1

28	38	30.9	39	4	US-09-101-059-11	Sequence 11, Appl1
29	37	30.1	31	4	US-09-308-345A-32	Sequence 32, Appl1
30	37	30.1	31	4	US-09-308-345A-33	Sequence 33, Appl1
31	36	29.3	16	2	US-08-676-279-6	Sequence 6, Appl1
32	36	29.3	17	2	US-08-676-279-12	Sequence 12, Appl1
33	36	29.3	17	4	US-08-951-034B-14	Sequence 14, Appl1
34	36	29.3	20	4	US-08-635-886C-54	Sequence 54, Appl1
35	36	29.3	20	4	US-08-974-690C-54	Sequence 47, Appl1
36	35.5	28.9	24	1	US-08-244-701B-47	Sequence 47, Appl1
37	35.5	28.9	24	4	US-09-076-721-47	Sequence 47, Appl1
38	35.5	28.9	33	4	US-08-469-260A-306	Sequence 306, App
39	35.5	28.9	33	4	US-08-488-446-306	Sequence 306, App
40	35.5	28.9	33	4	US-08-467-344A-306	Sequence 306, App
41	35	28.5	21	3	US-09-184-658-68	Sequence 68, Appl1
42	35	28.5	21	4	US-09-504-262D-68	Sequence 68, Appl1
43	35	28.5	37	4	US-09-314-268-109	Sequence 109, App
44	34	27.6	7	1	US-08-460-343B-6	Sequence 6, Appl1
45	34	27.6	7	1	US-08-398-028B-6	Sequence 6, Appl1

#### ALIGNMENTS

```
RESULT 1
US-10-133-007-5
; Sequence 5, Application US/10133007
; Patent No. 6623921
; GENERAL INFORMATION:
; APPLICANT: Aoyagi, Katsumi
; APPLICANT: Aoyagi, Katsumi
; APPLICANT: Iida, Kuniharu
; APPLICANT: Iida, Kuniharu
; APPLICANT: Yagi, Shinaro
; TITLE OF INVENTION: METHOD FOR MEASUREMENT OF HEPATITIS C VIRUS
; FILE REFERENCE: 594.352USWO
; CURRENT APPLICATION NUMBER: US/10/133,007
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US/09/509,449
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 10-216094
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/JP99/04129
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fused polypeptide including Hepatitis C virus sequence.
US-10-133-007-5

Query Match          100.0%; Score 123; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.0e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRHRSRVNG 21
Db 1 PRGSRPSWGPTDPRHRSRVNG 21

RESULT 2
PCT-US92-07865-15
; Sequence 15, Application PC/TUS9207865
; GENERAL INFORMATION:
; APPLICANT: Dreesman, Gordon R.
; APPLICANT: Burk, Kenneth H.
; APPLICANT: Pauletti, Daniel
; TITLE OF INVENTION: Peptide-based Hepatitis C Virus
; TITLE OF INVENTION: Immunoassays
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Peter Dehlinger
```

STREET: 350 Cambridge Avenue, Suite 300  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/07865  
FILING DATE: 19920916  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
FILING DATE: 16-SEP-1991  
FILING DATE: 12-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 1600-0086.41  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Hepatitis C Virus  
INDIVIDUAL ISOLATE: Dp10, amino acids 98-121 of the HCV  
INDIVIDUAL ISOLATE: polypotein  
PCT-US92-07865-15

Query Match 91.1%; Score 112; DB 5; Length 24;  
Best Local Similarity 90.5%; Pred. No. 1.1e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PRGSRPSWGPDPHRRSRNVG 21  
Db 3 PRGSRPSWGPDPHRRSRNVG 23

RESULT 3  
US-08-905-054B-3  
Sequence 3, Application US/08905054B  
Patent No. 6596476  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Lesniowski, Richard R.  
APPLICANT: Leung, Tat K.  
TITLE OF INVENTION: HEPATITIS C ASSAY  
FILE REFERENCE: 4767.US.C7  
CURRENT APPLICATION NUMBER: US/08/905.054B  
CURRENT FILING DATE: 1997-08-01  
PRIOR APPLICATION NUMBER: US 08/707,355  
PRIOR FILING DATE: 1996-09-04  
PRIOR APPLICATION NUMBER: US 08/507,740  
PRIOR FILING DATE: 1995-07-26  
PRIOR APPLICATION NUMBER: US 08/373,920  
PRIOR FILING DATE: 1995-01-17  
PRIOR APPLICATION NUMBER: US 08/183,207  
PRIOR FILING DATE: 1994-01-18  
PRIOR APPLICATION NUMBER: US 07/760,292  
PRIOR FILING DATE: 1991-09-16  
PRIOR APPLICATION NUMBER: US 07/610,180  
PRIOR FILING DATE: 1990-11-07  
PRIOR APPLICATION NUMBER: US 07/456,162

PRIOR FILING DATE: 1989-12-22  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Hepatitis C Virus  
US-08-905-054B-3

Query Match 91.1%; Score 112; DB 4; Length 28;  
Best Local Similarity 90.5%; Pred. No. 1.3e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PRGSRPSWGPDPHRRSRNVG 21  
Db 2 PRGSRPSWGPDPHRRSRNVG 22

RESULT 4  
PCT-US92-07813-3  
Sequence 3, Application PC/TUS9207813  
GENERAL INFORMATION:  
APPLICANT: LESNIEWSKI, RICHARD R.  
APPLICANT: LEUNG, TAT K.  
TITLE OF INVENTION: HEPATITIS C ASSAY  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESS: ABBOTT LABORATORIES  
STREET: ONE ABBOTT PARK ROAD  
CITY: ABBOTT PARK  
STATE: ILLINOIS  
COUNTRY: U.S.A.  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/07813  
FILING DATE: 19920916  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: FOREMSKIP, PRISCILLA E.  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 4767.P3.03  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-937-9556  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US92-07813-3

Query Match 91.1%; Score 112; DB 5; Length 28;  
Best Local Similarity 90.5%; Pred. No. 1.3e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PRGSRPSWGPDPHRRSRNVG 21  
Db 2 PRGSRPSWGPDPHRRSRNVG 22

RESULT 5  
US-07-681-701-10  
Sequence 10, Application US/07681701  
Patent No. 5574132  
GENERAL INFORMATION:  
APPLICANT: Lacroix, Martial

TITLE OF INVENTION: PEPTIDES AND MIXTURES THEREOF FOR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 875 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/681,701  
FILING DATE: 19910405  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: IAF-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 715-0742  
TELEFAX: (212) 715-0673  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-681-701-10

Query Match 81.3%; Score 100; DB 1; Length 27;  
Best Local Similarity 85.0%; Pred. No. 6.7e-08;  
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RGSRRPSWGPTDPRRRSNG 21  
Db 1 RGSRRPSWGPTDPRRRSNG 20

RESULT 6  
US-08-635-886C-8  
Sequence 8, Application US/08635886C  
Patent No. 6555114  
GENERAL INFORMATION:  
APPLICANT: LEROUX-ROELS, Geert  
APPLICANT: DELEYS, Robert  
APPLICANT: MAERTENS, Geert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
TITLE OF INVENTION: VIRUS  
FILE REFERENCE: 2752-18  
CURRENT APPLICATION NUMBER: US/08/635,886C  
CURRENT FILING DATE: 1996-04-25  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 8  
LENGTH: 20  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-08-635-886C-8

Query Match 76.4%; Score 94; DB 4; Length 20;  
Best Local Similarity 94.1%; Pred. No. 3.5e-07;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRRRS 17  
Db 4 PRGSRPSWGPTDPRRRS 20

RESULT 7  
US-08-974-690C-8  
Sequence 8, Application US/08974690C  
Patent No. 6613333  
GENERAL INFORMATION:  
APPLICANT: LEROUX-ROELS, Geert  
APPLICANT: DELEYS, Robert  
APPLICANT: MAERTENS, Geert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
TITLE OF INVENTION: VIRUS  
FILE REFERENCE: 2551-94  
CURRENT APPLICATION NUMBER: US/08/974,690C  
CURRENT FILING DATE: 1997-11-19  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 8  
LENGTH: 20  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-08-974-690C-8

Query Match 76.4%; Score 94; DB 4; Length 20;  
Best Local Similarity 94.1%; Pred. No. 3.5e-07;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRRRS 17  
Db 4 PRGSRPSWGPTDPRRRS 20

RESULT 8  
US-08-604-365-22  
Sequence 22, Application US/08604365  
Patent No. 6183949  
GENERAL INFORMATION:  
APPLICANT: Seidel, Christoph; Erlich-Weinreich,  
APPLICANT: Gertraud; Bayer, Hubert; Wiennues, Ursula; Jung,  
APPLICANT: Gether-Gernard; Ihlenfeldt, Hans Georg  
TITLE OF INVENTION: HCV Peptide Antigens and Methods for  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM PS/2 PC-DOS  
OPERATING SYSTEM: Wordperfect  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,365  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,398  
FILING DATE: 11-MARCH-1993  
APPLICATION NUMBER: PCT/EP92/01468  
FILING DATE: 30-JUNE-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 41 22 160.5  
FILING DATE: 04-JULY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 41 41 304.0  
FILING DATE: 14-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 42 09 216.9  
FILING DATE: 21-MARCH-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Tsai, Christine H.  
REGISTRATION NUMBER: 34,266  
REFERENCE/DOCKET NUMBER: BOER 1010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-604-365-22

Query Match 69.1%; Score 85; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.1e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPR 14  
DB 1 PRGSRPSWGPTDPR 14

RESULT 9  
US-09-689-678-22  
Sequence 22, Application US/09689678  
Patent No. 6592871  
GENERAL INFORMATION:  
APPLICANT: Seidel, Christoph, Ehrlich-Weinreich,  
Gertraud; Bayer, Hubert; Wienhues, Ursula; Jung,  
Gunter-Gerhard; Ihlenfeldt, Hans Georg  
TITLE OF INVENTION: HCV Peptide Antigens and Methods for  
the Determination of HCV  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/689, 678  
FILING DATE: 13-Oct-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,365  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 07/977, 398  
FILING DATE: 11-MARCH-1993  
APPLICATION NUMBER: FGT/EP92/01468  
FILING DATE: 30-JUNE-1992  
APPLICATION NUMBER: DE 41 22 160.5  
FILING DATE: 04-JULY-1991  
APPLICATION NUMBER: DE 41 41 304.0  
FILING DATE: 14-DEC-1991  
APPLICATION NUMBER: DE 42 09 216.9  
FILING DATE: 21-MARCH-1992  
ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.  
REGISTRATION NUMBER: 34,266  
REFERENCE/DOCKET NUMBER: BOER 1010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-689-678-22

Query Match 69.1%; Score 85; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.1e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPR 14  
DB 1 PRGSRPSWGPTDPR 14

RESULT 10  
US-08-473-475A-25  
Sequence 25, Application US/08473475A  
Patent No. 5843450  
GENERAL INFORMATION:  
APPLICANT: DAMSON, GEORGE  
APPLICANT: BRIDON, DOMINIQUE P.  
APPLICANT: SCHROEDER-POLIAK, PAMELLA A.  
APPLICANT: KNIGGE, MARK F.  
TITLE OF INVENTION: HEPATITIS GB VIRUS SYNTHETIC PEPTIDES  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ABBOTT LABORATORIES  
STREET: 100 ABBOTT PARK ROAD  
CITY: ABBOTT PARK  
STATE: IL  
COUNTRY: USA  
ZIP: 60064  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473, 475A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: POREMBSKI, PRISCILLA E.  
REGISTRATION NUMBER: 33206  
REFERENCE/DOCKET NUMBER: 5527US.P8  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-937-2623  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-473-475A-25

Query Match 52.0%; Score 64; DB 2; Length 31;  
Best Local Similarity 71.4%; Pred. No. 0.011;  
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 WGPDRHRSRNVG 21  
Db 4 WGRDPRHRSRNLG 17

RESULT 11  
US-08-635-886C-9  
Sequence 9, Application US/08635886C  
Patent No. 6553114  
GENERAL INFORMATION:  
APPLICANT: LEROUX-ROELS, Geert  
APPLICANT: DELEYS, Robert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
TITLE OF INVENTION: VIRUS  
FILE REFERENCE: 2752-18  
CURRENT APPLICATION NUMBER: US/08/635, 886C  
CURRENT FILING DATE: 1996-04-25  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 20  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-08-635-886C-9

Query Match 46.3%; Score 57; DB 4; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.07;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 PTPDRHRSRNVG 21  
Db 1 PTPDRHRSRNLG 12

RESULT 12  
US-08-974-690C-9  
Sequence 9, Application US/08974690C  
Patent No. 661333  
GENERAL INFORMATION:  
APPLICANT: LEROUX-ROELS, Geert  
APPLICANT: DELEYS, Robert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
TITLE OF INVENTION: VIRUS  
FILE REFERENCE: 2551-94  
CURRENT APPLICATION NUMBER: US/08/974, 690C  
CURRENT FILING DATE: 1997-11-19  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 20  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-08-974-690C-9

Query Match 46.3%; Score 57; DB 4; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.07;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 PTPDRHRSRNVG 21  
Db 1 PTPDRHRSRNLG 12

RESULT 13  
US-10-133-007-6  
Sequence 6, Application US/10133007  
Patent No. 6623921  
GENERAL INFORMATION:  
APPLICANT: Aoyagi, Katsumi  
APPLICANT: Ohue, Chiharu  
APPLICANT: Iida, Kumiko  
APPLICANT: Yagi, Shintaro  
TITLE OF INVENTION: METHOD FOR MEASUREMENT OF HEPATITIS C VIRUS  
FILE REFERENCE: 594.352USMO  
CURRENT APPLICATION NUMBER: US/10/133, 007  
CURRENT FILING DATE: 2002-04-26  
PRIOR APPLICATION NUMBER: US/09/509, 449  
PRIOR FILING DATE: 2000-03-28  
PRIOR APPLICATION NUMBER: JP 10-216094  
PRIOR FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: PCT/JP99/04129  
PRIOR FILING DATE: 1999-07-30  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6  
LENGTH: 20  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Fused polypeptide including Hepatitis C virus sequence.  
US-10-133-007-6

Query Match 45.5%; Score 56; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.098;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DPHRSRNVG 21  
Db 1 DPHRSRNVG 10

RESULT 14  
US-08-444-818-301  
Sequence 301, Application US/08444818  
Patent No. 6150087  
GENERAL INFORMATION:  
APPLICANT: Rutter, David Y.  
APPLICANT: Rutter, William J.  
TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444, 818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403, 590  
ATTORNEY/AGENT INFORMATION:  
FILING DATE: 14-MAR-1995  
NAME: Hardin, Ailsa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885

; INFORMATION FOR SEQ ID NO: 301:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-444-818-301

Query Match 40.7%; Score 50; DB 3; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GSGRPSWP 10  
 DB 1 GSGRPSWP 8

RESULT 15  
 US-08-444-818-299  
 ; Sequence 299, Application US/08444818  
 ; Patent No. 6150087  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chien, David Y.  
 ; APPLICANT: Rutter, William J.  
 ; TITLE OF INVENTION: NANBY Diagnostics and Vaccines  
 ; NUMBER OF SEQUENCES: 777  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Chiron Corporation  
 ; STREET: 4560 Horton Street  
 ; CITY: Emeryville  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94608-2916  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/444,818  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/403,590  
 ; FILING DATE: 14-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hardin, Alisa A.  
 ; REGISTRATION NUMBER: 33,895  
 ; REFERENCE/DOCKET NUMBER: 0110,002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (508)359-3876  
 ; TELEFAX: (508)359-3885  
 ; INFORMATION FOR SEQ ID NO: 299:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-444-818-299

Query Match 39.8%; Score 49; DB 3; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSW 8  
 DB 1 PRGSRPSW 8

Search completed: June 29, 2004, 12:09:59  
 Job time: 17.2941 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 12:08:25 ; Search time 43.2353 Seconds

(without alignments)  
137.314 Million cell updates/sec

Title: US-09-509-449C-5  
Perfect score: 123  
Sequence: 1 PRGSRPSCWPTDPRHRSRNG 21

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 309702

Minimum DB seq length: 0  
Maximum DB seq length: 40

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCF\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppa/PCFUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	91.1	25	10	US-09-891-983A-29
2	112	91.1	25	14	US-10-173-480-56
3	105	85.4	30	12	US-10-296-734-470
4	94	76.4	20	12	US-10-651-165-8
5	90	73.2	30	12	US-10-296-734-472
6	85	63.1	15	14	US-10-371-540-22
7	78	63.4	25	10	US-09-891-983A-30
8	78	63.4	25	14	US-10-173-480-57
9	67	54.5	15	14	US-10-268-561-11
10	67	54.5	15	14	US-10-268-561-11
11	65	52.8	15	14	US-10-268-561-12
12	65	52.8	15	14	US-10-268-561-12
13	57	46.3	20	14	US-10-268-561-12
14	54	43.9	18	15	US-10-350-719-180
15	49	39.8	9	8	US-08-344-824-83

16	49	39.8	17	10	US-09-891-983A-28	Sequence 28, Appl
17	49	39.8	17	14	US-10-173-480-55	Sequence 55, Appl
18	44.5	36.2	35	14	US-10-029-386-33805	Sequence 33805, A
19	44	35.8	29	10	US-09-933-767-630	Sequence 630, Appl
20	44	35.8	29	12	US-10-004-860-530	Sequence 630, Appl
21	44	35.8	29	14	US-10-023-282-530	Sequence 630, Appl
22	44	35.8	31	14	US-10-023-282-530	Sequence 630, Appl
23	44	35.8	31	14	US-10-023-282-530	Sequence 630, Appl
24	40.5	32.9	35	10	US-09-836-392-30	Sequence 28430, A
25	40.5	32.9	35	10	US-09-836-392-30	Sequence 28430, A
26	40.5	32.9	35	10	US-09-836-392-30	Sequence 28430, A
27	40	32.5	39	12	US-09-925-302-681	Sequence 681, Appl
28	40	32.5	10	14	US-10-228-806-61	Sequence 681, Appl
29	40	32.5	10	15	US-10-100-303A-48	Sequence 48, Appl
30	40	32.5	31	14	US-10-144-929-252	Sequence 252, Appl
31	39.5	32.1	37	14	US-10-144-929-252	Sequence 252, Appl
32	39.5	32.1	37	14	US-10-029-386-31485	Sequence 31485, A
33	39.5	32.1	40	14	US-10-008-524A-106	Sequence 106, Appl
34	39.5	32.1	40	15	US-10-350-719-106	Sequence 106, Appl
35	39	31.7	9	8	US-08-344-824-84	Sequence 84, Appl
36	39	31.7	9	16	US-10-440-390-29	Sequence 84, Appl
37	39	31.7	24	14	US-10-194-441A-83	Sequence 23, Appl
38	38.5	31.3	31	9	US-09-880-578-36	Sequence 83, Appl
39	38.5	31.3	35	9	US-09-864-761-37459	Sequence 37459, A
40	38.5	31.3	40	14	US-10-008-524A-107	Sequence 107, Appl
41	38.5	31.3	40	15	US-10-350-719-107	Sequence 107, Appl
42	38	30.9	21	14	US-10-084-813-347	Sequence 347, Appl
43	38	30.9	21	14	US-10-084-813-348	Sequence 348, Appl
44	38	30.9	21	14	US-10-084-813-349	Sequence 349, Appl
45	38	30.9	28	10	US-09-974-879-564	Sequence 564, Appl
					US-09-305-736-569	Sequence 569, Appl

#### ALIGNMENTS

RESULT 1  
US-09-891-983A-29  
Sequence 29, Application US/09891983A  
Publication No. US20030108858A1  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Shah, Dinesh O.  
APPLICANT: Dawson, George A.  
APPLICANT: Muerhoff, A. Scott  
APPLICANT: Jiang, Lily  
APPLICANT: Gutierrez, Robin A.  
APPLICANT: Leary, Thomas P.  
APPLICANT: Desai, Suresh  
APPLICANT: Stewart, James L.  
TITLE OF INVENTION: Methods For The simultaneous Detection  
TITLE OF INVENTION: OF HCV Antigens And HCV Antibodies  
FILE REFERENCE: 6821 US 01  
CURRENT APPLICATION NUMBER: US/09/891, 983A  
CURRENT FILING DATE: 2001-06-26  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 29  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Hepatitis C Virus  
US-09-891-983A-29

Query Match 91.1%, Score 112, DB 10, Length 25;  
Best Local Similarity 90.5%, Pred. No. 1.7e-07;  
Matches 19, Conservative 1, Mismatches 1, Gaps 0;

1 PRGSRPSCWPTDPRHRSRNG 21  
5 PRGSRPSCWPTDPRHRSRNG 25  
Db

RESULT 2  
US-10-173-480-56

Sequence 56, Application US/10173480  
Publication No. US20030152948A1  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Shah, Dinesh O.  
APPLICANT: Dawson, George A.  
APPLICANT: Muenhoff, A. Scott  
APPLICANT: Jiang, Lily  
APPLICANT: Gutierrez, Robin A.  
APPLICANT: Leary, Thomas P.  
APPLICANT: Desai, Suresh  
APPLICANT: Stewart, James L.  
TITLE OF INVENTION: METHODS FOR THE SIMULTANEOUS DETECTION  
OF HCV ANTIGENS AND HCV ANTIBODIES  
FILE REFERENCE: 6821 US P1  
CURRENT APPLICATION NUMBER: US/10/173,480  
CURRENT FILING DATE: 2002-06-17  
PRIOR APPLICATION NUMBER: 09/891,983  
PRIOR FILING DATE: 2001-06-26  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 56  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HCV-Core derived peptides  
US-10-173-480-56

Query Match 91.1% Score 112; DB 14; Length 25;  
Best Local Similarity 90.5%; Pred. No. 1.7e-07;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRHRSNNGV 21  
|||||  
DB 5 PRGSRPSWGPTDPRHRSNNGV 25

RESULT 3  
US-10-296-734-420  
Sequence 420, Application US/10296734  
Publication No. US20040054137A1  
GENERAL INFORMATION:  
APPLICANT: Thompson, Scott A  
APPLICANT: Ramshaw, Ian A  
TITLE OF INVENTION: Synthetic molecules and uses therefor  
FILE REFERENCE: Savine  
CURRENT APPLICATION NUMBER: US/10/296,734  
CURRENT FILING DATE: 2003-08-04  
PRIOR APPLICATION NUMBER: AU P07761/00  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 1507  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 420  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: HepC 1a segment 7  
US-10-296-734-420

Query Match 85.4% Score 105; DB 12; Length 30;  
Best Local Similarity 94.7%; Pred. No. 1.6e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRHRSN 19  
|||||  
DB 12 PRGSRPSWGPTDPRHRSN 30

RESULT 4  
US-10-651-165-8  
Sequence 8, Application US/10651165

Publication No. US20040047877A1  
GENERAL INFORMATION:  
APPLICANT: LEROUX-ROELS, Geert  
APPLICANT: DELEYS, Robert  
APPLICANT: MAERTENS, Geert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
FILE REFERENCE: 2551-94  
CURRENT APPLICATION NUMBER: US/10/651,165  
CURRENT FILING DATE: 2003-09-02  
PRIOR APPLICATION NUMBER: US/08/974,690C  
PRIOR FILING DATE: 1997-11-19  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 20  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-10-651-165-8

Query Match 76.4% Score 94; DB 12; Length 20;  
Best Local Similarity 94.1%; Pred. No. 2.5e-05;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPRHRS 17  
|||||  
DB 4 PRGSRPSWGPTDPRHRS 20

RESULT 5  
US-10-296-734-422  
Sequence 422, Application US/10296734  
Publication No. US20040054137A1  
GENERAL INFORMATION:  
APPLICANT: Thompson, Scott A  
APPLICANT: Ramshaw, Ian A  
TITLE OF INVENTION: Synthetic molecules and uses therefor  
FILE REFERENCE: Savine  
CURRENT APPLICATION NUMBER: US/10/296,734  
CURRENT FILING DATE: 2003-08-04  
PRIOR APPLICATION NUMBER: AU P07761/00  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 1507  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 422  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: HepC 1a segment 8  
US-10-296-734-422

Query Match 73.2% Score 90; DB 12; Length 30;  
Best Local Similarity 88.2%; Pred. No. 0.00013;  
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPSWGPTDPRHRSNNGV 21  
|||||  
DB 1 RPSWGPTDPRHRSNNGV 17

RESULT 6  
US-10-371-540-22  
Sequence 22, Application US/10371540  
Publication No. US20030198644A1  
GENERAL INFORMATION:  
APPLICANT: Roche Diagnostics GmbH  
TITLE OF INVENTION: HCV PEPTIDE ANTIGENS AND METHODS FOR THE DETERMINATION OF HCV  
FILE REFERENCE: 9793/129



```

; CURRENT APPLICATION NUMBER: US/10/371,540
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 09/689,678
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 08/604,365
; PRIOR FILING DATE: 1996-02-21
; PRIOR APPLICATION NUMBER: P 41 41 304.1
; PRIOR FILING DATE: 1991-12-14
; PRIOR APPLICATION NUMBER: P 42 09 215.9
; PRIOR FILING DATE: 1992-03-21
; PRIOR APPLICATION NUMBER: PCT/EP92/01468
; PRIOR FILING DATE: 1992-06-30
; PRIOR APPLICATION NUMBER: P 41 22 160.5
; PRIOR FILING DATE: 1991-07-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent version 3.1
; SEQ ID NO 22
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide antigen for anti-HCV antibodies
US-10-371-540-22

Query Match          69.1%; Score 85; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPTDPR 14
DB 1 PRGSRPSWGPTDPR 14

RESULT 7
US-09-891-983A-30
; Sequence 30, Application US/09891983A
; Publication No. US20030108858A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Shah, Dinesh O.
; APPLICANT: Dawson, George A.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Jiang, Lily
; APPLICANT: Gutierrez, Robin A.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Desai, Suresh
; APPLICANT: Stewart, James L.
; TITLE OF INVENTION: Methods For The simultaneous Detection
; FILE REFERENCE: 6821.US.01
; CURRENT APPLICATION NUMBER: US/09/891,983A
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-891-983A-30

Query Match          63.4%; Score 78; DB 10; Length 25;
Best Local Similarity 86.7%; Pred. No. 0.0038;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 SWGPTDPRRRSRNG 21
DB 1 SWGPTDPRRRSRNG 15

RESULT 8
US-10-173-480-57
; Sequence 57, Application US/10173480
; Publication No. US20030152948A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Shah, Dinesh O.
; APPLICANT: Dawson, George A.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Jiang, Lily
; APPLICANT: Gutierrez, Robin A.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Desai, Suresh
; APPLICANT: Stewart, James L.
; TITLE OF INVENTION: METHODS FOR THE SIMULTANEOUS DETECTION
; FILE REFERENCE: 6821.US.P1
; CURRENT APPLICATION NUMBER: US/10/173,480
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/891,983
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV-Core derived peptides
US-10-173-480-57

Query Match          63.4%; Score 78; DB 14; Length 25;
Best Local Similarity 86.7%; Pred. No. 0.0038;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 SWGPTDPRRRSRNG 21
DB 1 SWGPTDPRRRSRNG 15

RESULT 9
US-10-268-561-11
; Sequence 11, Application US/10268561
; Publication No. US20030148333A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies
; FILE REFERENCE: CDS0286
; CURRENT APPLICATION NUMBER: US/10/268,561
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/337453
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent version 3.1
; SEQ ID NO 11
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-561-11

Query Match          54.5%; Score 67; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPT 11
DB 5 PRGSRPSWGPT 15

RESULT 10
US-10-268-569-11
; Sequence 11, Application US/10268569
; Publication No. US20030152965A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: HCV Core Protein Sequences
; FILE REFERENCE: CDS-0286
```

CURRENT APPLICATION NUMBER: US/10/268,569  
CURRENT FILING DATE: 2002-10-10  
PRIOR APPLICATION NUMBER: 60/347,303  
PRIOR FILING DATE: 2001-11-11  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 11  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Hepatitis C virus  
US-10-268-569-11

Query Match 54.5%; Score 67; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.061;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGSRPSWGPT 11  
DB 5 PRGSRPSWGPT 15

RESULT 11  
US-10-268-561-12  
Sequence 12, Application US/10268561  
Publication No. US2003014833A1  
GENERAL INFORMATION:  
APPLICANT: Ortho-Clinical Diagnostics, Inc.  
TITLE OF INVENTION: HCV Anti-Core Monoclonal Antibodies  
FILE REFERENCE: CDS0286  
CURRENT APPLICATION NUMBER: US/10/268,561  
CURRENT FILING DATE: 2002-10-10  
PRIOR APPLICATION NUMBER: US 60/337453  
PRIOR FILING DATE: 2001-11-05  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Hepatitis C virus  
US-10-268-561-12

Query Match 52.8%; Score 65; DB 14; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 SWGPTDPRHRSR 18  
DB 1 SWGPTDPRHRSR 12

RESULT 12  
US-10-268-569-12  
Sequence 12, Application US/10268569  
Publication No. US20030152965A1  
GENERAL INFORMATION:  
APPLICANT: Ortho-Clinical Diagnostics, Inc.  
TITLE OF INVENTION: HCV Core Protein Sequences  
FILE REFERENCE: CDS 0288  
CURRENT APPLICATION NUMBER: US/10/268,569  
CURRENT FILING DATE: 2002-10-10  
PRIOR APPLICATION NUMBER: 60/347,303  
PRIOR FILING DATE: 2001-11-11  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 12  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Hepatitis C virus  
US-10-268-569-12

Query Match 52.8%; Score 65; DB 14; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 SWGPTDPRHRSR 18  
DB 1 SWGPTDPRHRSR 12

RESULT 13  
US-10-651-165-9  
Sequence 9, Application US/10651165  
Publication No. US20040047877A1  
GENERAL INFORMATION:  
APPLICANT: LEROUX-ROELS, Geert  
APPLICANT: DELBYS, Robert  
APPLICANT: MAERTENS, Geert  
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
TITLE OF INVENTION: VIRUS  
FILE REFERENCE: 2551-94  
CURRENT APPLICATION NUMBER: US/10/651,165  
CURRENT FILING DATE: 2003-09-02  
PRIOR APPLICATION NUMBER: US/08/974,690C  
PRIOR FILING DATE: 1997-11-19  
PRIOR APPLICATION NUMBER: PCT/EP94/03555  
PRIOR FILING DATE: 1994-10-28  
PRIOR APPLICATION NUMBER: EP 93402718.6  
PRIOR FILING DATE: 1993-11-04  
NUMBER OF SEQ ID NOS: 286  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 20  
TYPE: PRT  
ORGANISM: hepatitis C virus  
US-10-651-165-9

Query Match 46.3%; Score 57; DB 12; Length 20;  
Best Local Similarity 83.3%; Pred. No. 1.5;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 PTDPRHRSRNVG 21  
DB 1 PTDPRHRSRNVG 12

RESULT 14  
US-10-350-719-180  
Sequence 180, Application US/10350719  
Publication No. US20030219726A1  
GENERAL INFORMATION:  
APPLICANT: Dootbar, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
FILE REFERENCE: 18396/2162  
CURRENT APPLICATION NUMBER: US/10/350,719  
CURRENT FILING DATE: 2003-01-24  
PRIOR APPLICATION NUMBER: PCT/GB01/01176  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: GB0018140.4  
PRIOR FILING DATE: 2000-07-24  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 180  
LENGTH: 18  
TYPE: PRT  
ORGANISM: Human papillomavirus type 51  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (18)-(18)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-10-350-719-180

Query Match 43.9%; Score 54; DB 15; Length 18;  
Best Local Similarity 57.1%; Pred. No. 3.3;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Wed Jun 30 11:14:03 2004

us-09-509-449c-5.rapb

Page 5

Qy 6 PSMGPTDPRHRSRN 19  
Db 3 PAMAPKPKPRHNSN 16

RESULT 15  
US-08-344-824-83  
; Sequence 83, Application US/08344824  
; Publication No. US20030152580A1  
; GENERAL INFORMATION:  
; APPLICANT: SETTE, Alessandro  
; APPLICANT: SIDNEY, John  
; TITLE OF INVENTION: HLA BINDING PEPTIDES AND THEIR USES  
; NUMBER OF SEQUENCES: 399  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Stewart Street Tower, 20th  
; STREET: Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,824  
; FILING DATE: 23-NOV-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,634  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 14137-80-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 83:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-344-824-83

Query Match 39.8%; Score 49; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. NO. 1e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0;

Qy 1 PRGSRPSW 8  
Db 2 PRGSRPSW 9

Search completed: June 29, 2004, 12:21:22  
Job time : 44.2353 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 12:02:29 ; Search time 14 Seconds  
(without alignments)  
144.287 Million cell updates/sec

Title: US-09-509-449C-5  
Perfect score: 123  
Sequence: 1 PRGSRPWSGPTDPRHRSRVG 21

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 9205

Minimum DB seq length: 0  
Maximum DB seq length: 40

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	38	30.9	23	2	S10448
2	37.5	30.5	29	2	C40638
3	34	27.6	37	2	B70566
4	33.5	27.2	24	2	S43011
5	33	26.8	28	2	S51593
6	33	26.8	33	1	IRTR1A
7	33	26.8	37	2	B7154
8	32.5	26.4	29	2	T34643
9	32	26.0	38	2	S17889
10	32	26.0	38	2	D95027
11	32	26.0	38	2	D97898
12	32	26.0	38	2	S58601
13	31	25.2	33	2	C21211
14	31	25.2	33	2	E21211
15	31	25.2	33	2	A42592
16	31	25.2	38	2	G45095
17	31	25.2	39	2	S71186
18	30	24.4	11	2	S70338
19	30	24.4	32	1	BSPGNB
20	30	24.4	38	2	D84988
21	30	24.4	40	2	A42087
22	29	23.6	27	2	C44908
23	29	23.6	28	2	A50692
24	29	23.6	31	2	S14615
25	29	23.6	32	1	S10NA1
26	29	23.6	33	1	IRTR59
27	29	23.6	33	1	IRTR42
28	29	23.6	33	2	A21211
29	29	23.6	33	2	D21211

30	29	23.6	33	2	S00710	protamine CII - ch
31	29	23.6	33	2	T01070	protamine - raibo
32	29	23.6	34	2	A24174	thermomycollin (EC
33	29	23.6	35	2	JN0369	microbial serine p
34	29	23.6	37	1	R5NT36	ribosomal protein
35	29	23.6	37	2	S78391	ribosomal protein
36	29	23.6	38	2	H83113	50S ribosomal prot
37	29	23.6	39	2	C97513	hypothetical prote
38	28.5	23.2	36	2	G87666	hypothetical prote
39	28	22.8	10	2	S53789	neuropeptide pec-H
40	28	22.8	20	2	PC2084	serine proteinase
41	28	22.8	20	2	S23981	outer layer protei
42	28	22.8	27	2	A28391	hypothetical prote
43	28	22.8	29	2	S08555	ribosomal protein
44	28	22.8	30	2	S14062	hypothetical prote
45	28	22.8	30	2	E45095	photosystem I ligh

## ALIGNMENTS

RESULT 1  
S10448  
nucleocapsid protein - porcine transmissible gastroenteritis virus (fragment)  
C/Species: porcine transmissible gastroenteritis virus  
C/Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 20-Sep-1999  
C/Accession: S10448  
R/Page, K.W.; Britton, P.; Bourneil, M.E.G.  
submitted to the EMBL Data Library, March 1990  
A/Reference number: S10447  
A/Accession: S10448  
A/Molecule type: mRNA  
A/Residues: 1-23 <PAG>  
A/Cross-references: EMBL:X52157; NID:G59010; PIDN:CAA36410.1; PID:G59011  
C/Suprafamily: coronavirus nucleocapsid protein  
C/Keywords: nucleocapsid

Query Match 30.9%; Score 38; DB 2; Length 23;  
Best Local Similarity 38.9%; Pred. No. 46;  
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 RGSRPWSGPTDPRHRSRV 19  
DB 4 QGQVSWGDESTKIRGRS 21

RESULT 2  
C40638  
orf 3' of cycl - Rhodobacter sphaeroides (fragment)  
C/Species: Rhodobacter sphaeroides  
C/Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
R/Accession: C40638  
R/Rot, M.A.; Wittuhn, V.C.; Schlake, B.A.; Soranno, M.; All, A.; Donohue, T.J.  
J. Bacteriol. 175, 358-366, 1993  
A/Title: Genetic evidence for the role of isocytochrome c2 in photosynthetic growth of R

A/Reference number: A40638; MUID:93123153; PMID:8380401  
A/Accession: C40638  
A/Status: preliminary  
A/Molecule type: nucleic acid  
A/Residues: 1-29 <ROR>  
A/Note: sequence extracted from NCBI backbone (NCBI:122345, NCBI:123561)

Query Match 30.5%; Score 37.5; DB 2; Length 29;  
Best Local Similarity 41.7%; Pred. No. 68;  
Matches 10; Conservative 2; Mismatches 7; Indels 5; Gaps 1;

QY 2 RGSRPWSGPTDPRHRSRV 20  
DB 5 QGFRPRGATGDRHTLGLKDRSV 28

RESULT 3  
B70566

Probable ribosomal protein l36 - Mycobacterium tuberculosis (strain H37RV)  
 C/Species: Mycobacterium tuberculosis  
 C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C/Accession: B70566  
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A/Reference number: A70500; MUID:98295987; PMID:9634230  
 A/Accession: B70566  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-37 <COL>  
 A/Cross-references: GB:295390; GB:AL123456; NID:G3261766; PIDN:CAB08727.1; PID:G2104384  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Gene: rpmJ  
 C/Superfamily: Escherichia coli ribosomal protein l36

Query Match 27.6%; Score 34; DB 2; Length 37;  
 Best Local Similarity 62.5%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
 DB 28 SPDRKOR 35

RESULT 4  
 S43011  
 merd protein - Yersinia enterocolitica transposon TN3926 (fragment)  
 C/Species: Yersinia enterocolitica  
 C/Date: 06-Jan-1995 #sequence\_revision 26-Jul-1996 #text\_change 08-Oct-1999  
 C/Accession: S43011  
 R/Osbourn, S.E.V.; Turner, A.K.; Grinstead, J.  
 submitted to: The EMBL Data Library, March 1994  
 A/Description: The structure of the bacterial transposable element, TN3926.  
 A/Reference number: S43011  
 A/Accession: S43011  
 A/Molecule type: DNA  
 A/Residues: 1-24 <OSB>  
 A/Cross-references: EMBL:X78059; NID:G460067; PIDN:CA454977.1; PID:G460068  
 C/Genetics:  
 A/Gene: merD  
 A/Mobile element: transposon TN3926

Query Match 27.2%; Score 33.5; DB 2; Length 24;  
 Best Local Similarity 43.8%; Pred. No. 2e+02;  
 Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 2 RGRSPSWGPTDP-RHR 16  
 DB 3 RPSKCMWPCQPNRHK 18

RESULT 5  
 S51593  
 myrB protein - Micromonospora griseorubida (fragment)  
 C/Species: Micromonospora griseorubida  
 C/Date: 19-Mar-1997 #sequence\_revision 18-Jul-1997 #text\_change 05-Nov-1999  
 C/Accession: S51593  
 R/Inouye, M.; Takada, Y.; Muro, N.; Beppu, T.; Horinouchi, S.  
 Mol. Gen. Genet. 245, 456-464, 1994  
 A/Title: Characterization and expression of a P-450-like mycinamicin biosynthesis gene u  
 A/Reference number: S51593; MUID:95107242; PMID:7808395  
 A/Accession: S51593  
 A/Molecule type: DNA  
 A/Residues: 1-28 <INO>  
 A/Cross-references: EMBL:D16098; NID:G286050; PIDN:BA03671.1; PID:d1004186; PID:G829046  
 C/Genetics:  
 A/Gene: myrB

Query Match 26.8%; Score 33; DB 2; Length 28;  
 Best Local Similarity 52.4%; Pred. No. 2.8e+02;  
 Matches 11; Conservative 0; Mismatches 6; Indels 4; Gaps 2;

QY 1 PRGRSPSWGPTDPHRSRNVG 21  
 DB 11 PRGRS--GP--PRARRGG 27

RESULT 6  
 IRTR1A  
 protamine 1A - rainbow trout  
 N/Alternate names: iridine  
 C/Species: Oncorhynchus mykiss (rainbow trout)  
 C/Date: 12-Aug-1981 #sequence\_revision 12-Aug-1981 #text\_change 16-Feb-1997  
 C/Accession: A02670  
 R/Randy, T.; Watanabe, S.  
 Inc. J. Protein Res. 1, 221-224, 1969  
 A/Title: A new method for fractionation of protamines and the amino acid sequences of sa  
 A/Reference number: A91774; MUID:72030546; PMID:4553052  
 A/Accession: A02670  
 A/Molecule type: protein  
 A/Residues: 1-33 <AND>  
 C/Superfamily: protamine Y2  
 C/Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match 26.8%; Score 33; DB 1; Length 33;  
 Best Local Similarity 45.0%; Pred. No. 3.3e+02;  
 Matches 9; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 PRGRSPSWGPTDPHRSRNV 20  
 DB 1 PRARRSSRPVRRRRRRRV 20

RESULT 7  
 D87154  
 50S ribosomal protein l36 [imported] - Mycobacterium leprae  
 C/Species: Mycobacterium leprae  
 C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
 C/Accession: D87154  
 R/Cole, S.T.; Bigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
 R.; Davies, R.M.; Devlin, K.; Duthey, S.; Felwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
 A/Title: Massive gene decay in the leprosy bacillus.  
 A/Reference number: A86909; MUID:21128732; PMID:11234002  
 A/Accession: D87154  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-37 <STO>  
 A/Cross-references: GB:AL450380; NID:G13093616; PIDN:CAC30916.1; GSPDB:GN00147  
 C/Genetics:  
 A/Gene: rpmJ  
 C/Superfamily: Escherichia coli ribosomal protein l36

Query Match 26.8%; Score 33; DB 2; Length 37;  
 Best Local Similarity 71.4%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 DPRHRSR 18  
 DB 29 DPRHRSR 35

RESULT 8  
 T34643  
 hypochemical protein SC10H5.05 SC10H5.05 - Streptomyces coelicolor  
 C/Species: Streptomyces coelicolor  
 C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C/Accession: T34643



A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-38 <MAT>  
 A:Cross-references: EMBL:X86563; NID:g902200; PIRN:CA6035.1; PID:g902270  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast

Query Match 26.0%; Score 32; DB 2; Length 38;  
 Best Local Similarity 33.3%; Pred. No. 5.2e+02;  
 Matches 6; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 RGSRRPSMGPTDPRHRSR 19  
 |||||  
 Db 4 RLSTSTWRKKPPKPMASN 21

RESULT 13  
 C21211  
 protamine TP16 - rainbow trout  
 C:Species: Oncorhynchus mykiss (rainbow trout)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: C21211  
 R:Aiken, J.M.; McKenzie, D.; Zhao, H.Z.; States, J.C.; Dixon, G.H.  
 Nucleic Acids Res. 11, 4907-4922, 1983  
 A>Title: Sequence homologues in the protamine gene family of rainbow trout.  
 A:Reference number: A21211; MUID:83272939; PMID:6308564  
 A:Accession: C21211  
 A:Molecule type: DNA  
 A:Residues: 1-33 <ATK>  
 A:Cross-references: EMBL:X01597; NID:g64338; PIRN:CA25750.1; PID:g64339  
 A>Note: the authors translated the codon CGC for residue 29 as Gly  
 C:Superfamily: protamine Y2  
 C:Keywords: DNA binding; nucleus

Query Match 25.2%; Score 31; DB 2; Length 33;  
 Best Local Similarity 44.4%; Pred. No. 6.2e+02;  
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 PRGSRPSMGPTDPRHRSR 18  
 |||||  
 Db 2 PRRRSSSRPRRRRRRAR 19

RESULT 14  
 E21211  
 protamine TP21 - rainbow trout  
 N:Alternate names: protamine 2a  
 C:Species: Oncorhynchus mykiss (rainbow trout)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 01-Aug-1997  
 C:Accession: E21211; C24970  
 R:Aiken, J.M.; McKenzie, D.; Zhao, H.Z.; States, J.C.; Dixon, G.H.  
 Nucleic Acids Res. 11, 4907-4922, 1983  
 A>Title: Sequence homologues in the protamine gene family of rainbow trout.  
 A:Reference number: A21211; MUID:83272939; PMID:6308564  
 A:Accession: E21211  
 A:Molecule type: DNA  
 A:Residues: 1-33 <ATK>  
 A:Cross-references: EMBL:X01599  
 R:McKay, D.J.; Renaux, B.S.; Dixon, G.H.  
 Eur. J. Biochem. 158, 361-366, 1986  
 A>Title: Rainbow trout protamines. Amino acid sequences of six distinct proteins from a  
 A:Reference number: A91170; MUID:86274711; PMID:3755398  
 A:Accession: C24970  
 A:Molecule type: protein  
 A:Residues: 2-33 <MCK>  
 C:Superfamily: protamine Y2  
 C:Keywords: DNA binding; nucleus  
 F:2-33/Product: protamine TP21 #status experimental <MAT>

Query Match 25.2%; Score 31; DB 2; Length 33;  
 Best Local Similarity 44.4%; Pred. No. 6.2e+02;

Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 PRGSRPSMGPTDPRHRSR 18  
 |||||  
 Db 2 PRRRSSSRPRRRRRRAR 19

RESULT 15  
 A42592  
 precorrin-6x reductase (EC 1.-.-.-) - Pseudomonas sp. (strain SC510) (fragments)  
 C:Species: Pseudomonas sp.  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Mar-1996  
 C:Accession: A42592  
 R:Banche, P.; Thibaut, D.; Favechon, A.; Debussche, L.; Cameron, B.; Crouzet, J.  
 U. Bacteriol. 174, 1036-1042, 1992  
 A>Title: Precorrin-6x reductase from Pseudomonas denitrificans: purification and character  
 A:Reference number: A42592; MUID:92121090; PMID:1732193  
 A:Accession: A42592  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-33 <BLA>  
 C:Keywords: oxidoreductase

Query Match 25.2%; Score 31; DB 2; Length 33;  
 Best Local Similarity 62.5%; Pred. No. 6.2e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GSRRSMP 10  
 |||||  
 Db 15 GLRPEWVP 22

Search completed: June 29, 2004, 12:09:05  
 Job time : 15 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 11:58:39 ; Search time 9.47059 Seconds

(without alignment)  
115.460 Million cell updates/sec

Title: US-09-509-449c-5  
Perfect score: 123  
Sequence: 1 PRGSRPSPW3PTDPRHRSRNVG 21

Scoring table: B:OSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 3234

Minimum DB seq length: 0  
Maximum DB seq length: 40

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	30.1	38	1	RL36_STRPY
2	34	27.6	37	1	RL36_MYCTU
3	34	27.6	38	1	RL36_RALSO
4	33	26.8	38	1	PR12_ONCMY
5	33	26.8	37	1	RL36_MYCLE
6	32	26.0	38	1	RL36_LACTA
7	31	25.2	32	1	PR17_ONCMY
8	31	25.2	32	1	PR18_ONCMY
9	31	25.2	37	1	RL36_HAELN
10	31	25.2	37	1	RL36_PASMU
11	30.5	24.8	38	1	FBR_MERTP
12	30	24.4	32	1	NEUB_PIG
13	30	24.4	37	1	RL36_LEBIN
14	30	24.4	37	1	RL36_SYEEL
15	30	24.4	38	1	RL36_BUCAL
16	30	24.4	38	1	RL36_BUCAP
17	30	24.4	40	1	MRAL_USRMA
18	29	23.6	27	1	CH12_STROI
19	29	23.6	32	1	PRT1_ONCKE
20	29	23.6	32	1	PRT5_ONCMY
21	29	23.6	32	1	PRT6_ONCMY
22	29	23.6	32	1	PRT9_ONCMY
23	29	23.6	34	1	THEM_MALSU
24	29	23.6	37	1	RK36_ARATH
25	29	23.6	37	1	RK36_EPRIV
26	29	23.6	37	1	RK36_LOTUA
27	29	23.6	37	1	RK36_OENHO
28	29	23.6	38	1	RL36_PSEAR
29	29	23.6	38	1	RL36_PSEPK
30	29	23.6	39	1	ABAR_BOMPA
31	28	22.8	14	1	TAT_HV1W2
32	28	22.8	14	1	TAT_HV1Z8
33	28	22.8	20	1	VMO2_CHICK

34	28	22.8	29	1	RL15_HALCU	P05971 halobacteri
35	28	22.8	34	1	PRT_DICLA	Q9627 dicentrarch
36	28	22.8	37	1	RL36_BIFLO	Q89326 bifidobacte
37	28	22.8	37	1	RL36_NEIMA	Q9152 neisseria m
38	28	22.8	37	1	RL36_STRGO	Q86772 streptomyce
39	27	22.0	27	1	PR1B_ACIQU	P02323 actinomer g
40	27	22.0	28	1	PRAC_PSEPO	P20360 pseudochis
41	27	22.0	30	1	AATW_RABIT	P12345 oryctolagus
42	27	22.0	33	1	PR1B_MUGCE	P08130 mugil cepha
43	27	22.0	34	1	PRT1_SAROR	P25327 sarda orien
44	27	22.0	34	1	PRT1_SCOSC	P83364 scomber sco
45	27	22.0	34	1	PRT2_THUTH	P02322 thunnus thy

## ALIGNMENTS

RESULT 1	ID	RL36_STRPY	STANDARD;	PRT;	38 AA.
AC	Q9A1V2				
DT	28-FEB-2003 (rel. 41, Created)				
DT	28-FEB-2003 (rel. 41, Last sequence update)				
DT	15-MAR-2004 (rel. 43, Last annotation update)				
DE	50S ribosomal protein L36.				
GN	RPMU OR SPY0076 OR SPYM3_0063 OR SPS0065 OR SPY18_0076 OR GBS0081 OR SAG0081 OR SMU.2003A.				
OS	Streptococcus pyogenes, (serotype M3)				
OS	Streptococcus pyogenes, (serotype M18)				
OS	Streptococcus agalactiae (serotype Iii)				
OS	Streptococcus agalactiae (serotype V), and				
OC	Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1314, 198466, 186103, 216495, 216466, 1309;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=S.pyogenes; STRAIN=SF370 / ATCC 700294 / Serotype M1;				
RX	MEDLINE=21192684; PubMed=11296296;				
RA	Ferrelli J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Prienau C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Via H.G., Najat F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;				
RT	"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=S.pyogenes; STRAIN=MGA5315 / Serotype M3;				
RX	MEDLINE=22133808; PubMed=1212206;				
RA	Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S., Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D., Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M., Schlievert P.M., Musser J.M.;				
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus: phase-encoded toxins, the high-virulence phenotype, and clone emergence.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=S.pyogenes; STRAIN=SSI-1 / Serotype M3;				
RX	MEDLINE=22683781; PubMed=12799345;				
RA	Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y., Okazaki N., Kawabata S., Yamazaki K., Shida T., Yasunaga T., Hayashi H., Hattori M., Hamada S.;				
RT	"Genome sequence of an M3 strain of Streptococcus pyogenes reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution.";				
RL	Genome Res. 13:1042-1055(2003).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=S.pyogenes; STRAIN=MGA58322 / Serotype M18;				
RX	MEDLINE=21927593; PubMed=11917108;				

RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
 RA Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,  
 RA Silva L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
 RA Kapur V., Daly J.A., Yeasly L.G., Musser J.M.,  
 RT "Genome sequence and comparative microarray analysis of serotype M18  
 RT group A Streptococcus strains associated with acute rheumatic fever  
 RT outbreaks.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
 RN [5].  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.agalactiae; STRAIN=NEW316 / Serotype III;  
 RX MEDLINE=22242508; PubMed=12354221;  
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier P., Frangoul L.,  
 RA Mesdek T., Zouine M., Couve E., Lallouli L., Foyat C., Tijen-Cuot P.,  
 RA Kunst F.,  
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
 RT invasive neonatal disease.";  
 RT Mol. Microbiol. 45:1499-1513(2002).  
 RL [6].  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=S.agalactiae; STRAIN=2603 V/R / Serotype V;  
 RX MEDLINE=22222588; PubMed=12300547;  
 RA Tetteijn H., Masiham V., Cieslewicz M.J., Bisen J.A., Peterson S.,  
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,  
 RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac L.M., Daugherty S.C.,  
 RA Desoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,  
 RA Raeburn D., Fedorova N.B., Scanlan D., Khouri H., Mulligan M., Mora M.,  
 RA Carey H.A., Cline R.T., Van Aken S.E., Gill U., Scarselli M., Mora M.,  
 RA Jacobini E.T., Bretton C., Gall G., Martini M., Vagni F., Malone D.,  
 RA Rinaldo D., Rappaport R., Telford J.L., Kasper D.L., Grandi G.,  
 RA Fraser C.M.,  
 RT "Complete genome sequence and comparative genomic analysis of an  
 RT emerging human pathogen, serotype V Streptococcus agalactiae.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).  
 RL [7].  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=S.mutans; STRAIN=UA159 / ATCC 700610 / Serotype C;  
 RX MEDLINE=22295063; PubMed=12397186;  
 RA Adic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Khan S., Qian Y.,  
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.,  
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental  
 RT pathogen.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
 RL [8].  
 RN -1- SIMILARITY: Belongs to the 136p family of ribosomal proteins.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL, AB006478; AAK33205.1;  
 DR EMBL, AE014137; AAM78670.1;  
 DR EMBL, AP005141; BAC63160.1;  
 DR EMBL, AE009959; AAL96899.1;  
 DR EMBL, AL766843; CAD45726.1;  
 DR EMBL, AE014195; AAM98989.1;  
 DR EMBL, AE015023; AAM9607.1;  
 DR HSP, P80256; IDP.  
 DR Sagaliet, gds0081;  
 DR TIGR, SAG0081;  
 DR HAMAP, MF\_00251; 1.  
 DR InterPro, IPR000473; Ribosomal L36.  
 DR Pfam, PF00444; Ribosomal L36; 1.  
 DR ProDom, PD002101; Ribosomal L36; 1.  
 DR TIGRFAms, TIGR01022; rpsL\_bact; 1.  
 DR PROSITE, PS00828; RIBOSOMAL\_L36; 1.  
 KW Ribosomal protein, Complete proteome.  
 SQ SEQUENCE 38 AA; 4451 MW; 893FBE9A5939D8 CRC64;

Query Match 30.1%; Score 37; DB 1; Length 38;  
 Beel Local Similarity 55.6%; Pred. No. 44;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 10 PTPRRHR 18  
 Db 28 PTPRRHR 36

## RESULT 2

ID RL36 MYCTU STANDARD; PRT; 37 AA.

AC P45810; O08382;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 50S ribosomal protein L36 (Ribosomal protein B).  
 GN RPLM OR RY3461C OR MT3567.1 OR MTCY13E12.14C OR MB3490C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OK NCBI\_Taxid=1773, 1765;  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Mole S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.,  
 RT "Complete genome sequence of Mycobacterium tuberculosis from the  
 RT Nature 393:537-544 (1998).  
 RL [2].  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Enomura M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Minkula A.,  
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.,  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RT J. Bacteriol. 184:5479-5490(2002).  
 RN [3].  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.bovis; STRAIN=BCG;  
 RX MEDLINE=96200849; PubMed=8621083;  
 RA Dubnau E., Soares S., Huang T.J., Jacobs W.R. Jr.,  
 RT "Overproduction of mycobacterial ribosomal protein S13 induces  
 RT catalase/peroxidase activity and hypersensitivity to isoniazid in  
 RT Mycobacterium smegmatis.";  
 RT Gene 170:17-22(1996).  
 RL [4].  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=M.bovis; STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12788972;  
 RA Garnier T., Eigemeier K., Camus J.-C., Medina N., Mansoor H.,  
 RA Payot M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
 RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.,  
 RT "The complete genome sequence of Mycobacterium bovis.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
 RL [5].  
 RN -1- SIMILARITY: Belongs to the 136p family of ribosomal proteins.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----  
DR EMBL; Z95390; CAB08727.1; -  
DR EMBL; AE007160; AAK47907.1; -  
DR EMBL; U15140; AAB17596.1; -  
DR EMBL; BX248346; CAD95677.1; -  
DR PIR; B70566; B70566.  
DR HSSP; P80256; 1DPE.  
DR TIGR; MT3567.1; -  
DR TubercuList; RV3461c; -  
DR HAMAP; MF\_00251; -; 1.  
DR InterPro; IPR000473; Ribosomal\_L36.  
DR Pfam; PF00444; Ribosomal\_L36; 1.  
DR ProDom; PD002101; Ribosomal\_L36; 1.  
DR TIGRFAMs; TIGR01022; rpm1\_bact; 1.  
DR PROSITE; PS00828; RIBOSOMAL\_L36; 1.  
KW Ribosomal protein, Complete proteome.  
SQ SEQUENCE 37 AA; 4343 MW; C92A4845315F930F CRC64;

Query Match 27.6%; Score 34; DB 1; Length 37;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
DB 28 SDRHKKR 35

RESULT 3  
ID R136\_RALSO STANDARD; PRT; 38 AA.

AC O8XV34;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L36.  
GN RPM1 OR RSC2997 OR R501115.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GM11000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Brottier P., Camus J.C., Cartolico L.,  
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Signer P., Thebaud P., Whalen M., Wincker P., Levy M.,  
RA Weisenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RT Nature 415:497-502 (2002).  
CC -1- SIMILARITY: Belongs to the L36P family of ribosomal proteins.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----  
DR EMBL; AL646073; CAD16706.1; -  
DR HAMAP; MF\_00251; -; 1.  
DR InterPro; IPR000473; Ribosomal\_L36.  
DR Pfam; PF00444; Ribosomal\_L36; 1.  
DR ProDom; PD002101; Ribosomal\_L36; 1.  
DR TIGRFAMs; TIGR01022; rpm1\_bact; 1.

DR PROSITE; PS00828; RIBOSOMAL\_L36; 1.  
KW Ribosomal protein, Complete proteome.  
SQ SEQUENCE 38 AA; 4392 MW; 25F89F24F1065455 CRC64;

Query Match 27.6%; Score 34; DB 1; Length 38;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
DB 29 SDRHKKR 36

RESULT 4  
ID PR12\_ONCMY STANDARD; PRT; 33 AA.

AC P02328;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Protamine 1A (Iridine 1A).  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=72030546; PubMed=4953052;  
RT Ando T., Watanabe S.;  
RT "A new method for fractionation of protamines and the amino acid  
RT sequences of saline and three components of Iridine";  
RL Int. J. Protein Res. 1:221-224 (1969).  
CC -1- FUNCTION: Protamines substitute for histones in the chromatin of  
CC sperm during the haploid phase of spermatogenesis. They compact  
CC sperm DNA into a highly condensed, stable and inactive complex.  
CC -1- SUBCELLULAR LOCATION: Nucleus.  
CC -1- TISSUE SPECIFICITY: Testis.  
DR PIR; A02670; IRTRIA.  
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
KW Testis; DNA condensation; Nuclear protein.  
SQ SEQUENCE 33 AA; 4406 MW; F82C0A9C80A1FFP2 CRC64;

Query Match 26.8%; Score 33; DB 1; Length 33;  
Best Local Similarity 45.0%; Pred. No. 1.4e+02;  
Matches 9; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 PRGSRPWSGTPDPHRSRV 20  
DB 1 PRRRSSRPVRRRRRRRV 20

RESULT 5  
ID R136\_MYCLE STANDARD; PRT; 37 AA.

AC O9X7A2;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 50S ribosomal protein L36.  
GN RPM1 OR ML1961 OR MLCB1222.31C.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
NCBI\_TaxID=1169;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TN;  
RX MEDLINE=21128732; PubMed=11234002;  
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier C., Harris D.,  
RA Mungall K., Bauman D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,

RA Holroyd S., Hornby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrett B.G.;  
 RT "Massive gene decay in the leprosy bacillus.";  
 RL Nature 409:1007-1011(2001).  
 CC -1- SIMILARITY: Belongs to the 136p family of ribosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: A1049491; CAB39837.1; -;  
 DR EMBL: A1583923; CAC30916.1; -;  
 DR FIR: D87154; D87154.  
 DR HSSP: P80256; IDPE.  
 DR Leptoma; ML1961; -;  
 DR HAMAP: MF\_00251; -; 1.  
 DR InterPro: IPR000473; Ribosomal\_L36.  
 DR Pfam: PF00444; Ribosomal\_L36; 1.  
 DR TIGR: TIGR01022; rpsL\_bact; 1.  
 DR PROSITE: PS00828; RIBOSOMAL\_L36; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 37 AA; 4458 MW; D22PFA941710B496F CRC64;  
 CY 12 DRRHRSR 18  
 DB 29 DRRHRSR 35  
 RESULT 6  
 ID R136 LACIA STANDARD; PRT; 36 AA.  
 AC P27146;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L36.  
 GN RPSL OR L12071 OR SPO233 OR SPO232.  
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus lactis),  
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris),  
 OS Streptococcus pneumoniae, and  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 OX NCBI\_TaxID=1360, 1359, 1313, 171101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=L.l.cremoris; STRAIN=IL1403;  
 RX MEDLINE=21235186; PubMed=11337471;  
 RA Bolotin A., Mincker P., Mauger S., Jallou O., Malarme K.,  
 RA Weissenbach V., Shilich S.D., Sorokin A.;  
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
 RT lactis ssp. lactis IL1403.";  
 RL Genome Res. 11:731-753(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=L.l.cremoris; STRAIN=MG1614;  
 RX MEDLINE=92148378; PubMed=1783905;  
 RA Koyula T., Hemilae H.;  
 RT "Nucleotide sequence of a Lactococcus lactis gene cluster encoding  
 RT adenylate kinase, initiation factor 1 and ribosomal proteins.";  
 RL J. Gen. Microbiol. 137:2595-2600(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC SPECIES=L.l.cremoris; STRAIN=MG1614;  
 RX MEDLINE=91348245; PubMed=1908794;  
 RA Koyula T., Palva I., Hemilae H.;  
 RT "Nucleotide sequence of the secY gene from Lactococcus lactis and  
 RT identification of conserved regions by comparison of four secY  
 RT proteins.";  
 RL FEMS Lett. 288:114-118(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.pneumoniae; STRAIN=ATCC BAA-255 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916;  
 RA Tetteil H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S., Heidelberg J., Debay R.T., Haft D.H., Dodson R.U.,  
 RA Durkin A.S., Gwinn M., Kolony J.F., Nelson W.C., Peterson J.D.,  
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Holtzapple E., Khouri H., Wolf A.M., Ulfteback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Anguilo S., Dickinson T., Hickey E.K.,  
 RA Holt I.E., Loftus B.J., Yang F., Stuth H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.,  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 RT pneumoniae.";  
 RL Science 293:498-506(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.pneumoniae; STRAIN=ATCC BAA-255 / R6;  
 RX MEDLINE=2428245; PubMed=11544234;  
 RA Hoskins U., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,  
 RA Dehoff B.S., Estrem S.T., Filtz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Leibman R., Glas J.S., Khoja H., Kraft A.R., Lagace R.E.,  
 RA Gilmour D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,  
 RA McAlaren S.M., McHenry M., McLeester K., Mundy C.W., Nicas T.I.,  
 RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rokey P.,  
 RA Sun P.-W., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Scharnd P.L.,  
 RA Glas U.I.;  
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";  
 RL J. Bacteriol. 183:5709-5717(2001).  
 CC -1- SIMILARITY: Belongs to the 136p family of ribosomal proteins.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE006436; AAK06169.1; -;  
 DR EMBL: X59250; CAA1942.1; -;  
 DR EMBL: AE007336; AAK74413.1; -;  
 DR EMBL: AE008402; AAK99016.1; -;  
 DR PIR: D95027; D95027.  
 DR PIR: D97898; D97898.  
 DR PIR: S17989; S17989.  
 DR HSSP: P80256; IDPE.  
 DR TIGR: SPO233; -;  
 DR HAMAP: MF\_00251; -; 1.  
 DR InterPro: IPR000473; Ribosomal\_L36.  
 DR Pfam: PF00444; Ribosomal\_L36; 1.  
 DR ProDom: PD002101; Ribosomal\_L36; 1.  
 DR TIGR: TIGR01022; rpsL\_bact; 1.  
 DR PROSITE: PS00828; RIBOSOMAL\_L36; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 36 AA; 4421 MW; 892FEFE9A5939D88 CRC64;  
 CY 10 PTDPRHRSR 18  
 DB 28 PANDPKHRSR 36  
 Query Match 26.8%; Score 32; DB 1; Length 38;  
 Best Local Similarity 44.4%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

RESULT 7
PRT7 ONCMY STANDARD; PRT; 32 AA.
AC P08146;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Protamine TP16.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Procatanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_Taxid=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83272939; PubMed=6308564;
RA Aiken J.N., McKenzie D., Zhao H.-Z., States J.C., Dixon G.H.;
RT "Sequence homologues in the protamine gene family of rainbow trout."
RL Nucleic Acids Res. 11:4907-4922(1983).
CC -!- FUNCTION: Protamines substitute for histones in the chromatin of
CC sperm during the haploid phase of spermatogenesis. They compact
CC sperm DNA into a highly condensed, stable and inactive complex.
CC -!- SUBCELLULAR LOCATION: Nucleus.
CC -!- TISSUE SPECIFICITY: Testis.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X01597; CAA25750.1; -
DR PIR; C12111; C21211.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nucleic protein.
FT INIT MET
SQ SEQUENCE 32 AA; 4323 MW; B07A9C9D8CC796FF CRC64;
Query Match 25.2%; Score 31; DB 1; Length 32;
Best Local Similarity 44.4%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
QY 1 PRGSRPSWGPDPHRRSR 18
DB 1 FRRRSSRPVRRRRAR 18
RESULT 8
PRT8 ONCMY STANDARD; PRT; 32 AA.
AC P12817;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Protamine 2A.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Procatanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_Taxid=8022;
RN [1]
RP SEQUENCE.
RX MEDLINE=86274711; PubMed=3155398;
RA McKay D.J., Renaux B.S., Dixon G.H.;
RT "Rainbow trout protamines. Amino acid sequences of six distinct
RT proteins from a single testis."
RT Eur. J. Biochem. 158:361-366(1986).
CC -!- FUNCTION: Protamines substitute for histones in the chromatin of
CC sperm during the haploid phase of spermatogenesis. They compact
CC sperm DNA into a highly condensed, stable and inactive complex.

```

```

CC -!- SUBCELLULAR LOCATION: Nucleus.
CC -!- TISSUE SPECIFICITY: Testis.
-----
DR PIR; E21211; E21211.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nucleic protein.
SQ SEQUENCE 32 AA; 4224 MW; B07A9C9D90B796FF CRC64;
Query Match 25.2%; Score 31; DB 1; Length 32;
Best Local Similarity 44.4%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
QY 1 PRGSRPSWGPDPHRRSR 18
DB 1 FRRRSSRPVRRRRAR 18
RESULT 9
PRT9 ONCMY STANDARD; PRT; 37 AA.
AC P46361;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L36.
GN RPL36 OR RPL36 OR H10798.1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleisemann R.D., Adams W.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitznugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips R., Spriggs T., Hedblom E., Cotton M.D.,
RA Utecherback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Georghen N.S.M.,
RA Genem C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RT Science 269:496-512(1995).
CC -!- SIMILARITY: Belongs to the L36P family of ribosomal proteins.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U32762; AAC22457.1; -
DR HSP; P80256; IDPE.
DR TIGR; H10798.1; -
DR HAMAP; MF_00251; -; 1.
DR InterPro; IPR000473; Ribosomal L36.
DR Pfam; PF00444; Ribosomal L36; 1.
DR Prodom; PD002101; Ribosomal L36; 1.
DR TIGRFAMs; TIGR01022; rplL36; 1.
DR PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 37 AA; 4297 MW; 53AE5B5D828F83D CRC64;
Query Match 25.2%; Score 31; DB 1; Length 37;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 TDPHRRSR 18

```

Db 28 SDPKRQR 35

RESULT 10

ID RL36\_PASMU STANDARD; PRT; 37 AA.

AC P57912;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 50S ribosomal protein L36.

GN RPMJ OR RPL36 OR PM1394.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.

OC NCBI\_TaxID=747;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Pm70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.U., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida Pm70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -1- SIMILARITY: Belongs to the L36P family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE006177; AK03478.1; -.

DR HSSP; P80256; 1DEP.

DR HAMAP; MF\_00251; -; 1.

DR InterPro; IPR000473; Ribosomal\_L36.

DR Pfam; PF00444; Ribosomal\_L36; 1.

DR ProDom; PD002101; Ribosomal\_L36; 1.

DR TIGRfam; TIGR01022; rplL\_bact; 1.

DR PROSITE; PS00628; RIBOSOMAL\_L36; 1.

KW Ribosomal protein; Complete proteome.

SQ SEQUENCE 37 AA; 4279 MW; 53AE5B4CB39F83D CRC64;

Query Match 25.2%; Score 31; DB 1; Length 37;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18

DB 28 SDPKRQR 35

RESULT 11

ID FER\_METPR STANDARD; PRT; 38 AA.

AC P81542;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE zinc-containing ferredoxin (Seven-iron ferredoxin) (Fragment).

GN ZFX.

OS Metallophera prunae.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Metallophera.

OC NCBI\_TaxID=47304;

RN (1)

RP SEQUENCE.

RA Gomes C.M., Faria A., Carita J., Mendes J.C., Regalia M., Chicau P.,

RA Huber H., Stetter K.O., Teixeira M.;

RT "Di-cluster, seven iron ferredoxins from hyperthermophilic

RT Sulfolobales.";

RL J. Biol. Inorg. Chem. 3:499-507(1998).

CC -1- FUNCTION: Ferredoxins are iron-sulfur proteins that transfer

CC electrons in a wide variety of metabolic reactions.

CC -1- COFACTOR: Binds one 3Fe-4S and one 4Fe-4S cluster and one zinc

CC atom (By similarity).

CC -1- SIMILARITY: Belongs to the bacterial-type ferredoxin family.

DR HSSP; P55907; 1YXR.

KW Electron transport; Iron-sulfur; Repeat; 4Fe-4S; 3Fe-4S; Methylation;

KW Zinc.

FT DOMAIN 1 38 N-TERMINAL EXTENSION.

FT METAL 16 16 ZINC.

FT METAL 19 19 ZINC.

FT METAL 37 37 ZINC.

FT MOD RES 30 30 METHYLATION.

FT NON TER 38 38

SQ SEQUENCE 38 AA; 4006 MW; 256D378618BC9ED5 CRC64;

Query Match 24.8%; Score 30.5; DB 1; Length 38;

Best Local Similarity 31.0%; Pred. No. 3.5e+02;

Matches 9; Conservative 3; Mismatches 8; Indels 9; Gaps 1;

QY 2 RGRPSWG-----PTDPRHSRNVG 21

DB 7 RTRRPVGDHAGKRVAPADPPVKKALG 35

RESULT 12

ID NEUB\_PIG STANDARD; PRT; 32 AA.

AC P01297;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Neuromedin B-32 [Contains: Neuromedin B].

GN NMB.

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN (1)

RP SEQUENCE.

RX MEDLINE=85279454; PubMed=4026853;

RA Minamoto N., Stroh T., Kangawa K., Matsuo H.;

RT "Neuromedin B-32 and B-30: two 'big' neuromedin B identified in

RT porcine brain and spinal cord.";

RT Biochem. Biophys. Res. Commun. 130:685-691(1985).

RN (2)

RP SEQUENCE.

RA Minamoto N., Kangawa K., Matsuo H.;

RT "Neuromedin B and neuromedin C: two mammalian bombesin-like peptides

RT identified in pig spinal cord and brain.";

RT Regul. Pept. 19:127-127(1987).

RN (3)

RP SEQUENCE OF 23-32.

RX MEDLINE=83282813; PubMed=6882442;

RA Minamoto N., Kangawa K., Matsuo H.;

RT "Neuromedin B: a novel bombesin-like peptide identified in porcine

RT spinal cord.";

RT Biochem. Biophys. Res. Commun. 114:541-548(1983).

CC -1- FUNCTION: Stimulates smooth muscle contraction in a manner similar

CC to that of bombesin.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: Belongs to the bombesin/neuromedin B/zanarensin

CC family.

DR PIR; B60301; BSRGN.

DR InterPro; IPR000874; Bombesin.

DR Pfam; PF02044; Bombesin; 1.

DR PROSITE; PS00257; BOMBESIN; 1.

KW Bombesin family; Amidation.

FT PEPTIDE 23 32 NEUROMEDIN B.

FT MOD RES 32 32 AMIDATION.

SQ SEQUENCE 32 AA; 3654 MW; 7B3FC92C953DBE0 CRC64;

Query Match 24.4%; Score 30; DB 1; Length 32;

Best Local Similarity 35.7%; Pred. No. 3.4e+02;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 7 SWGPTDRHRSRV 20  
Db 4 SWDLPEPRSRAGK 17

## RESULT 13

RL36 LEPIIN

ID

RL36 LEPIIN

STANDARD;

PRT;

37 AA.

AC

Q8XD13;

DT

30-MAY-2000 (Rel. 39, Created)

DT

30-MAY-2000 (Rel. 39, Last sequence update)

DT

10-OCT-2003 (Rel. 42, Last annotation update)

DE

50S ribosomal protein L36.

GN

RPMJ OR LA0761.1.

OS

Leptospira interrogans.

OC

Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.

NX

NCBI\_TaxID=173;

[1]

SEQUENCE FROM N.A.

RC

STRAIN=Jai / Serogroup Icterohaemorrhagiae / Serovar Jai;

RX

MEDLINE=20088935; PubMed=10620683;

RA

Zuerner R.L., Hartkeerl R.A., van de Kemp H., Bal A.E.;

RT

"Characterization of the Leptospira interrogans S10-spc-alpha

operon".

RL

FEMS Microbiol. Lett. 182:303-308 (2000).

RN

[2]

SEQUENCE FROM N.A.

RC

STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar Jai;

RX

MEDLINE=2258143; PubMed=12712204;

RA

Ren S.-X., Fu G., Qiang X.-G., Zeng R., Miao Y.-G., Xu H.,

RT

Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,

RN

Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,

RC

Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,

RX

Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,

RA

Saint Girons I., Somerville R.L., Wen Y.-M., Shi W.-H., Chen Z.,

RT

Xu J.-G., Zhao G.-P.;

RN

Unique physiological and pathogenic features of Leptospira

RC

interrogans revealed by whole-genome sequencing.";

RX

Nature 422:888-893 (2003).

RA

-1- SIMILARITY: Belongs to the L36P family of ribosomal proteins.

RT

This SWISS-PROT entry is copyright. It is produced through a collaboration

RN

between the Swiss Institute of Bioinformatics and the EMBL outstation -

RC

the European Bioinformatics Institute. There are no restrictions on its

RX

use by non-profit institutions as long as its content is in no way

RA

modified and this statement is not removed. Usage by and for commercial

RT

entities requires a license agreement (See http://www.isb-sib.ch/announce/

or send an email to license@isb-sib.ch).

RN

EMBL, AF115283; AAD40606.1; -

RC

EMBL, AE011262; -; NOT\_ANNOTATED\_CDS.

RX

EMBL, AE011263; -; NOT\_ANNOTATED\_CDS.

RA

HSSP, P80256; 1DEF.

RT

HAMAP, MP\_00251; -; 1.

RN

InterPro; IPR000473; Ribosomal\_L36.

RC

Pfam; PF00444; Ribosomal\_L36.

RX

ProDom; PD002101; Ribosomal\_L36.

RA

TIGRFAMs; TIGR01022; rpmJ\_bact; 1.

RT

PROSITE; PS00828; RIBOSOMAL\_L36; 1.

RN

Kw Ribosomal protein; Complete proteome.

RC

FT

CONFLICT 24 24 G -&gt; R (IN REF. 2)

RX

SEQUENCE 37 AA; 4182 MW; SFC68666BFF5D6 CRC64;

RA

Query Match

Best Local Similarity 24.4%; Score 30; DB 1; Length 37;

RX

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

RA

Qy 11 TDPHRSR 18

Db 28 TNPGRKOR 35

ID

RL36 BUCAI

AC

P57570;

DT

16-OCT-2001 (Rel. 40, Created)

DT

16-OCT-2001 (Rel. 40, Last sequence update)

DT

28-FEB-2003 (Rel. 41, Last annotation update)

DE

50S ribosomal protein L36.

GN

RPMJ OR BUS03.

OS

Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

OC

synbiotic bacterium).

NX

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

RT

Enterobacteriaceae; Buchnera.

RN

NCBI\_TaxID=118099;

RC

[1]

SEQUENCE FROM N.A.

RX

STRAIN=Tokyo 1998;

RA

MEDLINE=20445173; PubMed=10993077;

RT

Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RN

Genome sequence of the endocellular bacterial symbiont of aphids

```

RT Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -1- SIMILARITY: belongs to the L36P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL/ AP001119; BAB13196.1; -.
DR HSSP; P80256; 1DPE.
DR HAMAP; MF_00251; ?; 1.
DR InterPro; IPR000473; Ribosomal_L36.
DR Pfam; PF00444; Ribosomal_L36; 1.
DR ProDom; PD002101; Ribosomal_L36; 1.
DR TIGRFAMs; TIGR01022; rpm_dact; 1.
DR PROSITE; PS00828; RIBOSOMAL_L36; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 38 AA; 4365 MW; 29F271414AA25621 CRC64;

```

```

Query March 24.4%; Score 30; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 4,1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 12 DRRHRSR 18
   |||:|
Db 30 DPKHKQR 36

```

Search completed: June 29, 2004, 12:06:32  
 Job time : 10.4706 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 29, 2004, 12:01:59 ; Search time 38.2941 Seconds  
(without alignments)  
173.026 Million cell updates/sec

```
Title: US-09-509-449C-5
Perfect score: 123
Sequence: 1 PRGSRPSWGPTDPRHRSRVG 21
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 31556

```
Minimum DB seq length: 0
Maximum DB seq length: 40
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```

1:  sp. archaea: *
2:  sp. bacteria: *
3:  sp. fungi: *
4:  sp. human: *
5:  sp. invertebrate: *
6:  sp. mammal: *
7:  sp. mtc: *
8:  sp. organelle: *
9:  sp. phage: *
10: sp. plant: *
11: sp. rodent: *
12: sp. virus: *
13: sp. vertebrate: *
14: sp. unclassified: *
15: sp. virus: *
16: sp. bacteriag: *
17: sp. archaeag: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	40.5	32.9	39	2	Q9X689	Q9X689 mercury res
2	39.5	32.1	33	6	Q77492	Q77492 nlyclitebus
3	38	30.9	21	11	Q92249	Q92249 mus musculus
4	38	30.9	23	12	Q66184	Q66184 porcine resein
5	38	30.9	23	12	Q66204	Q66204 transmissisella
6	38	30.9	37	16	Q7WR41	Q7WR41 bordetella
7	38	30.9	37	16	Q7W2D3	Q7W2D3 bordetella
8	38	30.9	37	16	Q7VTA9	Q7VTA9 bordetella
9	38	30.9	40	15	Q9YFG8	Q9YFG8 rous sarcom
10	37.5	30.5	29	2	Q0S844	Q0S844 rhodobacter
11	36	29.3	29	5	Q9SWF4	Q9SWF4 acropora mi
12	36	29.3	39	12	Q66546	Q66546 human herpe
13	35	28.5	34	5	Q819U9	Q819U9 periplaneta
14	35	28.5	37	16	Q7YKFA	Q7YKFA naemophilu
15	34.5	28.0	22	4	Q9S597	Q9S597 homo sapien
16	34.5	28.0	35	13	Q801Y0	Q801Y0 carassius a

17	33.5	27.2	24	13	056907	Yersinia en
18	33	26.8	20	1.3	042534	042534 Brachydeni
19	33	26.8	29	15	091004	091004 human immu
20	33	26.8	29	15	091002	091002 human immu
21	33	26.8	29	15	091009	091009 human immu
22	33	26.8	29	15	0910J3	0910J3 human immu
23	33	26.8	35	16	070008	070008 rhodospirell
24	33	26.8	38	16	0700M5	0700M5 rhodospirell
25	33	26.8	40	2	09R4G1	09R4G1 bacillus sp
26	32.5	26.4	29	16	086497	086497 streptomyc
27	32	26.0	25	5	018630	018630 sphaearechin
28	32	26.0	26	12	056480	056480 hepatitis c
29	32	26.0	26	13	056481	056481 hepatitis c
30	32	26.0	28	13	042517	042517 myxine glut
31	32	26.0	29	15	091004	091004 human immu
32	32	26.0	29	15	0910J7	0910J7 human immu
33	32	26.0	29	15	0910J5	0910J5 human immu
34	32	26.0	29	15	0910J6	0910J6 human immu
35	32	26.0	38	8	033303	033303 zea mays (m
36	32	26.0	38	15	088384	088384 simian t-Ly
37	32	26.0	38	16	0833E1	0833E1 enterococci
38	32	26.0	39	4	0900E6	0900E6 homo sapien
39	32	26.0	39	5	018620	018620 biopharaxi
40	32	26.0	39	16	0700E9	0700E9 rhodospirell
41	32	26.0	39	16	070050	070050 rhodospirell
42	31.5	25.6	36	16	0700J4	0700J4 rhodospirell
43	31	25.2	12	12	066002	066002 transmissib
44	31	25.2	17	12	085004	085004 porcine rees
45	31	25.2	27	4	0900A6	0900A6 homo sapien

## ALIGNMENTS

```

RESULT 1
Q3x689
ID Q3x689 PRELIMINARY; PRT; 39 AA.
AC Q3x689;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MeaD (Fragment).
GN MeaD.
OS mercury resistant bacterium '96 SE13.
OC Bacteria.
CX NCBI_Taxid=93601;
RV [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96 SE13;
RA Holt R.J., Bruce K.D., Strike P.;
RT "Conservation of transposon structures in soil bacteria.";
RL FEMS Microbiol. Ecol. 0:0-0(1999).
DR EMBL; AF134211; AAD34398.1; -.
FT NON_TER
FT SEQUENCE 39 AA; 4421 MW; 9A35BCB0F0F5D258 CRC64;
SQ

Query Match 33.9%; Score 40.5; DB 2; Length 39;
Best Local Similarity 40.0%; Pred. No. 77;
Matches 10; Conservative 3; Mismatches 5; Indels 7; Gaps 2.

QY 3 GSRP-----SWGPTDP-RHRSRVV 20
      ||| ||| ||| |||
      :|:|:|:|:|:|:|:|:|
      5 GARPWPASKCMWPQCQPNRHSTRV 29

RESULT 2
O77492
ID O77492 PRELIMINARY; PRT; 33 AA.
AC O77492;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE D4 dopamine receptor (D4DR) (Fragment).

```

OS Nycitebus coucang (Slow loris).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Strepsirrhini; Loridae; Nycitebus.  
 OX NCBI\_TaxID=470;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Inoue-Murayama M., Takenaka O., Murayama Y.,  
 RT "Origin and divergence of tandem repeats of primate D4 dopamine  
 receptor genes.";  
 RL Primates 39:217-224(1998).  
 DR EMBL: AB016201; BAA32039.1; -;  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 KW Receptor.  
 FT NON TER 1  
 FT NON TER 33  
 SQ SEQUENCE 33 AA; 3541 MW; 4AF17501A4FD5CEC CRC64;

Query Match 32.1%; Score 39.5; DB 6; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 91;  
 Matches 12; Conservative 1; Mismatches 4; Indels 5; Gaps 2;

QY 1 RGSRRPSWGPTDPRH---NSR 18  
 DB 1 PR-SRPAPGDLPPRRLLGPRSR 21

RESULT 3  
 ID Q922Y9 PRELIMINARY; PRT; 21 AA.  
 AC Q922Y9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC006692; AA006692.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 21 AA; 2315 MW; 60698D703D0C1692 CRC64;

Query Match 30.9%; Score 38; DB 11; Length 21;  
 Best Local Similarity 75.0%; Pred. NO. 94;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RGSRRPSW 8  
 DB 14 RGSRRPSW 21

RESULT 4  
 ID Q66184 PRELIMINARY; PRT; 23 AA.  
 AC Q66184;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Coronavirus PRCV nucleoprotein, 5' region (fragment).  
 OS Porcine respiratory coronavirus.  
 OC Coronavirus; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11146;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=PRCV 86/137004(British Isolate);  
 RC Page K.W., Britton P., Boursnell M.E.G.;  
 RT "Sequence analysis of the leader RNA of two porcine coronaviruses:  
 transmissible gastroenteritis virus and porcine respiratory

RT coronavirus";  
 RL Virus Genes 0:0-0(1990).  
 DR EMBL: X52668; CA936895.1; -;  
 DR PIR: S10448; S10448.  
 DR GO: GO:0019013; C:Viral nucleocapsid; IEA.  
 KW Nucleocapsid.  
 FT NON TER 23  
 FT NON TER 23  
 SQ SEQUENCE 23 AA; 2565 MW; D761FC279767540F CRC64;

Query Match 30.9%; Score 38; DB 12; Length 23;  
 Best Local Similarity 38.9%; Pred. NO. 1e+02;  
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 RGSRRPSWGPTDPRHRSN 19  
 DB 4 QGQRVSWGDESTKIRGRS 21

RESULT 5  
 ID Q66204 PRELIMINARY; PRT; 23 AA.  
 AC Q66204;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Coronavirus TGEV nucleoprotein, 5' region (fragment).  
 OS Transmissible gastroenteritis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11149;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=TGEV F872/70 (Virulent British Isolate);  
 RC Page K.W., Britton P., Boursnell M.E.G.;  
 RT "Sequence analysis of the leader RNA of two porcine coronaviruses:  
 transmissible gastroenteritis virus and porcine respiratory  
 coronavirus";  
 RL Virus Genes 0:0-0(1990).  
 DR EMBL: X52157; CA936410.1; -;  
 DR PIR: S10448; S10448.  
 DR GO: GO:0019013; C:Viral nucleocapsid; IEA.  
 KW Nucleocapsid.  
 FT NON TER 23  
 FT NON TER 23  
 SQ SEQUENCE 23 AA; 2565 MW; D761FC279767540F CRC64;

Query Match 30.9%; Score 38; DB 12; Length 23;  
 Best Local Similarity 38.9%; Pred. NO. 1e+02;  
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 RGSRRPSWGPTDPRHRSN 19  
 DB 4 QGQRVSWGDESTKIRGRS 21

RESULT 6  
 ID Q7MRAL PRELIMINARY; PRT; 37 AA.  
 AC Q7MRAL;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE 50S ribosomal protein 136.  
 GN RPM1 OR B80053.  
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=RB50 / ATCC BAA-588;  
 RC MEDLINE=22827654; PubMed=12910271;  
 RX Parkhill U., Sebahia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Umwin L., Whitehead S., Barrett B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,  
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL: BX640437; CAE30555.1;  
 KM Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 37 AA; 4332 MW; 53B3BBE5487DADD CRC64;

Query Match 30.9%; Score 38; DB 16; Length 37;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
 |||||:  
 Db 28 TDPHRSR 35

RESULT 7  
 Q7WZD3 PRELIMINARY; PRT; 37 AA.  
 AC Q7WZD3;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 50S ribosomal protein 136.  
 GN RPMJ OR BP0053.  
 OS *Bordetella parapertussis*.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=12822 / ATCC BAA-587;  
 RX MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Umwin L., Whitehead S., Barrett B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,  
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL: BX640423; CAE39794.1;  
 KM Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 37 AA; 4332 MW; 53B3BBE5487DADD CRC64;

Query Match 30.9%; Score 38; DB 16; Length 37;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
 |||||:  
 Db 28 TDPHRSR 35

RESULT 8  
 Q7VTA9 PRELIMINARY; PRT; 37 AA.  
 AC Q7VTA9;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 50S ribosomal protein 136.  
 GN RPMJ OR BP3638.  
 OS *Bordetella pertussis*.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 12251;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Umwin L., Whitehead S., Barrett B.G., Maskell D.J.,  
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,  
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL: BX640422; CAE3895.1;  
 KM Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 37 AA; 4332 MW; 53B3BBE5487DADD CRC64;

Query Match 30.9%; Score 38; DB 16; Length 37;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
 |||||:  
 Db 28 TDPHRSR 35

RESULT 9  
 Q9YFG8 PRELIMINARY; PRT; 40 AA.  
 AC Q9YFG8;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE (Prague "wild type") src gene; amino acids 1-40 (fragment).  
 OS Rous sarcoma virus.  
 OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.  
 OC NCBI\_TaxID=11886;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8321633; PubMed=6304353;  
 RA Kitamura N., Yoshida M.,  
 RT "Small deletion in src of Rous sarcoma virus modifying transformation  
 RT phenotypes: Identification of 207-nucleotide deletion and its smaller  
 RT product with protein kinase activity,"  
 RL J. Virol. 46:985-992(1983).  
 DR EMBL: J02349; AAA42577.1;  
 FT NON TER 40  
 SQ SEQUENCE 40 AA; 3989 MW; 11DEC1BF9FBD3064 CRC64;

Query Match 30.9%; Score 38; DB 15; Length 40;  
 Best Local Similarity 53.8%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 GPTDPRHRSRNG 21  
 |||||:  
 Db 7 GPTDPRHRSRNG 19

RESULT 10  
 Q05846 PRELIMINARY; PRT; 29 AA.  
 ID Q05846  
 AC Q05846;

```

DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Isocytichrome c2 (cycl) (Fragment).
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RA Rott M.A., Withnum V.C., Schilke B.A., Soriano M., Ali A.,
RA Donohue T.J.;
RT "Genetic evidence for the role of isocytichrome c2 in photosynthetic
RT growth of Rhodobacter sphaeroides."
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL, U02104; AAA61342.1; -.
DR PIR, C40638; C40638.
FT NON TER 29
SQ SEQUENCE 29 AA; 3263 MW; E8461B87691D3278 CRC64;

Query Match 30.5%; Score 37.5; DB 2; Length 29;
Best Local Similarity 41.7%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 5; Gaps 1;

QY 2 RGSRPWGPTDPRH-----RSRV 20
Db 5 QGFRPRHGATGDRHTLGLKDRSV 28

RESULT 11
Q95WF4 PRELIMINARY; PRT; 27 AA.
ID Q95WF4
AC Q95WF4;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Nuclear receptor ANNR10 (Fragment).
OS Acropora millepora.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Astrocoelina; Acroporidae; Acropora.
NCBI_TaxID=45264;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21490824; PubMed=11603940;
RA Grasso L.C., Hayward D.C., Trueman J.W.H., Hardie K.M., Janssens P.A.,
RA Ball E.E.;
RT "The evolution of nuclear receptors: evidence from the coral
RT Acropora."
RL MOL. Phylogenet. Evol. 21:93-102(2001).
CC -1 SIMILARITY BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -1 SIMILARITY BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR EMBL, AF323690; AAU29203.1; -.
DR GO, GO:0005634; C:nucleus; IEA.
DR GO, GO:0004872; F:receptor activity; IEA.
DR GO, GO:0003700; F:transcription factor activity; IEA.
DR GO, GO:0006355; P:regulation of transcription; DNA-dependent; IEA.
DR GO, GO:0006350; P:transcription; IEA.
DR InterPro, IPR01628; Znf_C4steroid.
DR Pfam, PF00105; Zf-C4; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT NON TER 1
FT NON TER 27
SQ SEQUENCE 27 AA; 3302 MW; 2E0C29C3A9691CA5 CRC64;

Query Match 29.3%; Score 36; DB 5; Length 27;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 PTDPHRRR 18
Db 19 PVDRKRRR 27

```

```

RESULT 12
Q6546 PRELIMINARY; PRT; 39 AA.
ID Q6546
AC Q6546;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-NOV-1996 (TRENBLREL. 19, Last annotation update)
DE EBV B95-8 Cl(e) DNA with antigen coding ORF (Fragment).
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
NCBI_TaxID=10376;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88217505; PubMed=2835748;
RA Walls D., Perricaudet M., Gannon F.;
RT "The analysis of EBV proteins which are antigenic in vivo."
RT Nucleic Acids Res. 16:2859-2872(1988).
DR EMBL, X07530; CAA30405.1; -.
FT NON TER 39
SQ SEQUENCE 39 AA; 4280 MW; 7937120F24AF5774 CRC64;

Query Match 29.3%; Score 36; DB 12; Length 39;
Best Local Similarity 40.0%; Pred. No. 3.5e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 4; Gaps 1;

QY 1 PRGSRPSW---GPTDPRR 16
Db 16 PKAGCRSPASTGPSTPRSR 35

RESULT 13
Q819U9 PRELIMINARY; PRT; 34 AA.
ID Q819U9
AC Q819U9;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Large conductance calcium activated potassium channel psiO spliceform
DE 1C (Fragment).
NCBI_TaxID=6978;
RN [1]
RP SEQUENCE FROM N.A.
RA Deret C., Mesutut S., Walther C., Wicher D.;
RT "Cloning and Alternative Splicing of a large conductance calcium-
RT activated potassium channel from Periplaneta americana."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF452166; AAN76821.1; -.
DR GO, GO:0005216; F:ion channel activity; IEA.
KW Ionic channel.
FT NON TER 1
FT NON TER 34
SQ SEQUENCE 34 AA; 3753 MW; AEDA93254993EBBF CRC64;

Query Match 28.5%; Score 35; DB 5; Length 34;
Best Local Similarity 43.8%; Pred. No. 4.2e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 GSRPSWGPTDPRR 18
Db 17 GTRPYGGTLKRRR 32

RESULT 14
Q7VKF4 PRELIMINARY; PRT; 37 AA.
ID Q7VKF4
AC Q7VKF4;

```

DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 50S ribosomal protein L36.  
 GN RPM2 OR HD1955.  
 OS Haemophilus ducreyi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OC NCBI\_TaxID=730;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=35000HP / ATCC 700724;  
 RA Munson R.S., Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,  
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.,  
 RT "The complete genome sequence of Haemophilus ducreyi."  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE017156; AAP6675.1; -  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 37 AA; 4279 MW; 53A83EB09A2F94ED CRC64;

Query Match 28.5%; Score 35; DB 16; Length 37;  
 Best Local Similarity 62.5%; Pred. No. 4.6e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 TDPHRSR 18  
 |||:|:  
 Db 28 TDPKHQR 35

RESULT 15  
 ID 095597 PRELIMINARY; PRT; 22 AA.  
 AC 095597;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE Bcl-2-JH protein (Fragment).  
 GN BCL-2-JH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95316863; PubMed=7796416;  
 RA Ji W., Qn G.Z., Ye P., Zhang X.Y., Halabi S., Ehrlich M.,  
 RT "Frequent detection of bcl-2/JH translocations in human blood and  
 RT organ samples by a quantitative polymerase chain reaction assay."  
 RL Cancer Res. 55:2876-2882(1995).  
 DR EMBL; S78526; AAD14286.1; -  
 FT NON TER 1 1  
 SQ SEQUENCE 22 AA; 2440 MW; 39988F2A20589300 CRC64;

Query Match 28.0%; Score 34.5; DB 4; Length 22;  
 Best Local Similarity 58.3%; Pred. No. 3.2e+02;  
 Matches 7; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 5 RPSMGPTDPRHR 16  
 |||:|:  
 Db 10 RP-WGQGNRGR 20

Search completed: June 29, 2004, 12:08:20  
 Job time : 40.2941 secs

**This Page Blank (uspto)**